

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 31, 2004, 02:50:53 ; Search time 1614.64 Seconds  
(without alignments)  
722.185 Million cell updates/sec

Title: US-09-974-974-1  
Perfect score: 32  
Sequence: 1 ggcuccgcccgaugagagcagcucuc 32

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 238946

Minimum DB seq length: 0  
Maximum DB seq length: 60

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	15.8	49.4	55	9	BX655160	Arabidops
C 2	15.4	48.1	43	1	AA947665	AA947665 oc32402.s
C 3	15.2	47.5	48	1	AA861970	AA861970 oi45d12.s
C 4	15.2	47.5	57	9	DR1M24T	AL742081 Danio rer
C 5	15.2	47.5	59	8	BZ770511	BZ770511 SALK 1434
C 6	15.2	46.9	58	8	BH770391	BH770391 LLMGtag16
C 7	14.8	46.2	57	9	CR110068	CR110068 Forward s
C 8	14.8	46.2	60	6	CD948943	CD948943 SHA 46 Ge
C 9	14.6	45.6	32	8	AZ650179	AZ650179 IMR520P11
C 10	14.6	45.6	42	7	CF920754	CF920754 gmrhww3-
C 11	14.6	45.6	46	1	AV833578	AV833578 AV833578
C 12	14.6	45.6	52	2	AW246952	AW246952 2822591.5
C 13	14.6	45.6	52	9	CR144482	CR144482 Forward s
C 14	14.6	45.6	57	8	AZ819538	AZ819538 ZMO091A14
C 15	14.4	45.0	46	1	AA181388	AA181388 zp42f11.s
C 16	14.4	45.0	46	9	CR398085	CR398085 Arabidops
C 17	14.4	45.0	51	2	BF308407	BF308407 601890722
C 18	14.4	45.0	52	1	AA173932	AA173932 zp01e02.s
C 19	14.4	45.0	54	2	AW781356	AW781356 sk68922.y
C 20	14.4	45.0	55	2	BE748176	BE748176 601571564
C 21	14.2	44.4	42	1	AJ789426	AJ789426 AJ789426
C 22	14.2	44.4	50	1	AU107710	AU107710 AU107710
C 23	14.2	44.4	52	2	BE321553	BE321553 NF024G09I
C 24	14.2	44.4	54	7	CK348290	CK348290 LecFAFLP0

ALIGNMENTS

RESULT 1	25	14.2	44.4	58	9	AL753395	AL753395 Arabidops
BX655160/c	C 26	14	43.8	39	7	R42069	R42069 YF98c06.sl
LOCUS	C 27	14	43.8	43	1	AI647472	AI647472 uk42e12.x
DEFINITION	C 28	14	43.8	53	1	AA846708	AA846708 aj41b12.s
ACCESSION	C 29	14	43.8	53	6	CA339521	CA339521 NISC 1Y03
VERSION	C 30	14	43.8	56	8	BH911451	BH911451 SALK_0687
KEYWORDS	C 31	14	43.8	58	8	AZ792055	AZ792055 ZMO043E15
SOURCE	C 32	14	43.8	58	9	CG723363	CG723363 1119076A0
ORGANISM	C 33	13.8	43.1	27	8	AZ993846	AZ993846 ZMO279K05
	C 34	13.8	43.1	49	1	AI869979	AI869979 wl63d04.x
	C 35	13.8	43.1	55	4	BG759188	BG759188 602710848
	C 36	13.8	43.1	55	8	AZ918644	AZ918644 1006004H0
	C 37	13.8	43.1	58	1	AI039811	AI039811 ox49d04.x
	C 38	13.8	43.1	59	1	AI988246	AI988246 sc98a02.y
	C 39	13.8	43.1	59	5	BP134619	BP134619 BP134619
	C 40	13.8	43.1	59	7	CN933300	CN933300 000510AOF
	C 41	13.6	42.5	40	6	CA586986	CA586986 LBG26p15
	C 42	13.6	42.5	45	7	TI7565	TI7565 mps v253 Th
	C 43	13.6	42.5	49	1	AA691190	AA691190 vt34c05.r
	C 44	13.6	42.5	52	1	AA674719	AA674719 vm74c06.s
	C 45	13.6	42.5	54	8	BZ289822	BZ289822 SALK_0232

55 bp DNA linear GSS 04-APR-2004  
Arabidopsis thaliana T-DNA flanking sequence GK-587C03-021291,  
genomic survey sequence.

BX655160  
BX655160.1 GI:37611548  
GSS.  
Arabidopsis thaliana (thale cress)

Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1  
Li, Y., Rosso, M.G., Strizhov, N., Viehoveer, P., and Weisshaar, B.  
GABI-Kat SimpleSearch: a flanking sequence tag (FST) database for  
the identification of T-DNA insertion mutants in Arabidopsis  
thaliana  
Bioinformatics 19 (11), 1441-1442 (2003)

JOURNAL  
MEDLINE  
PUBMED  
22755829  
12874060

REFERENCE  
AUTHORS  
TITLE  
Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and  
Weisshaar, B.  
An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for  
flanking sequence tag-based reverse genetics  
Plant Mol. Biol. 53 (1-2), 247-259 (2003)

JOURNAL  
MEDLINE  
PUBMED  
23117147  
14756321

REFERENCE  
AUTHORS  
TITLE  
Strizhov, N., Li, Y., Rosso, M.G., Viehoveer, P., Dekker, K.A. and  
Weisshaar, B.  
High-throughput generation of sequence indexes from T-DNA  
mutagenized Arabidopsis thaliana lines  
Biotechniques 35 (6), 1164-1168 (2003)

JOURNAL  
PUBMED  
14682050  
4 (bases 1 to 55)

REFERENCE  
AUTHORS  
TITLE  
Li, Y., Strizhov, N., Rosso, M.G. and Weisshaar, B.  
Direct Submission  
Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer  
Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany  
This sequence has been recovered from the left border of the T-DNA.  
It indicates an insertion within the locus defined by BAC clone  
K18L8. Details on the protocols used for generation of the sequence  
are described in References 1-3. The sequences are generated at the  
MPI for Plant Breeding Research in the context of the GABI-Kat  
project. GABI-Kat is part of the German Plant Genomics program

designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.  
 Location/Qualifiers  
 1..55

# FEATURES

source

/organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /strain="Columbia 0"  
 /db\_xref="taxon:3702"  
 /clone="GK-587C03-021291"  
 /clone\_lib="Arabidopsis thaliana T-DNA insertion lines"  
 /ecotype="Col-0"  
 /note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pGAB11 (GenBank accession number: AY529716). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

## ORIGIN

Query Match 49.4%; Score 15.8; DB 9; Length 55;  
 Best Local Similarity 51.9%; Pred. No. 6.5e+04;  
 Matches 14; Conservative 6; Mismatches 7; Indels 0; Gaps 0;  
 Qy 6 UGCGCUGAGAGUGAGUCUCUUC 32  
 Db 47 TCGACTGATGAGCGAGGAATCTTC 21

## RESULT 2

AA947665  
 LOCUS  
 DEFINITION oq32d02.s1 NCI\_CGAP GC4 Homo sapiens cDNA clone IMAGE:1588035 3', similar to TR:Q13539 Q13539 MARINER TRANSPOSASE. ;, mRNA sequence.  
 AA947665  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 43)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-t@mail.nih.gov](mailto:cgapbs-t@mail.nih.gov)  
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Trace considered overall poor quality  
 Insert Length: 1878 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amersham  
 High quality sequence stop: 1.  
 Location/Qualifiers

## FEATURES

source

1..43  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1588035"  
 /tissue\_type="pooled germ cell tumors"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI\_CGAP\_GC4"  
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from 3 pooled germ cell tumors, and was then primed with a Not I -

oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 48.1%; Score 15.4; DB 1; Length 43;  
 Best Local Similarity 76.5%; Pred. No. 9.2e+04;  
 Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 11 UGAUGAGAGUGAGG 27

Db 3 TGATGAGGATGAGC 19

## RESULT 3

AA861970/c  
 LOCUS  
 DEFINITION o145d12.s1 NCI\_CGAP HN3 Homo sapiens cDNA clone IMAGE:1485623 3', similar to gb:Z19574\_rnal KERATIN, TYPE I CYTOSKELETAL 17 (HUMAN);, mRNA sequence.  
 AA861970  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 48)  
 NCI/NIDR-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute / National Institute of Dental Research, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-t@mail.nih.gov](mailto:cgapbs-t@mail.nih.gov)  
 Tissue Procurement: John Ensley, M.D.  
 CDNA Library Preparation: Stratagene, Inc.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)

Trace considered overall poor quality  
 Insert Length: 1607 Std Error: 0.00  
 Seq primer: -40m13 fwd. ET from Amersham  
 High quality sequence stop: 1.  
 Location/Qualifiers

1..48  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1485623"  
 /tissue\_type="squamous cell carcinoma from base of tongue"  
 /lab\_host="NCI\_CGAP HN3"  
 /clone\_lib="NCI\_CGAP HN3"  
 /note="Organ: tongue; Vector: Bluescript SK-; Site: 1; EcoRI; Site: 2; XhoI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.0 kb. 5' adaptor sequence: 5' GAATTCGGCAGCAG 3' 3' adaptor sequence: 5' (GA)10ACTAGTCGAGTTTTTTTTTTTTTTT 3'"

## ORIGIN

Query Match 47.5%; Score 15.2; DB 1; Length 48;  
 Best Local Similarity 46.4%; Pred. No. 1.1e+05;  
 Matches 13; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy 4 CCUGGCCUGAGUGAGUGAGCUCUU 31

Db 36 CTTTGCCTGATGACATAAAGCTCTGTT 9

Result No.	Query Match	Score	Query			DB	ID	Description
			Length	Length	No.			
1	32	100.0	32	6	AA140446		AA140446	Maxizyme-
2	21.4	66.9	33	6	AA140448		AA140448	Maxizyme-
3	17.8	55.6	60	6	ABN44085		ABN44085	Human spl
4	17.4	54.4	29	2	AAA14986		AAA14986	Maxizyme
5	17.4	54.4	40	6	AA140455		AA140455	Aryl hydr
6	16.6	51.9	30	2	AAZ32140		AAZ32140	Human PRO
7	16.6	51.9	30	3	AAZ89576		AAZ89576	Human PRO
8	16.6	51.9	30	3	AAZ89588		AAZ89588	Human PRO
9	16.6	51.9	30	3	AAA46937		AAA46937	Probe use
10	16.6	51.9	30	10	ADJ58635		ADJ58635	Human PRO
C 11	16.4	51.2	27	6	ABSG6252		ABSG6252	Anti-huma
12	16.4	51.2	45	2	AAI96613		AAI96613	Cyclitol-
13	16.4	51.2	55	4	AAI14065		AAI14065	Vector pA
14	16.2	50.6	39	2	AAV54276		AAV54276	Bacteriop
15	16.2	50.6	39	2	AAV53358		AAV53358	Soluble s
C 16	16.2	50.6	47	3	AAZ69149		AAZ69149	Human map
C 17	16.2	50.6	51	5	AAAF8327		AAAF8327	RAR-gamma
C 18	16.2	50.6	60	6	ABN47364		ABN47364	Human spl
19	16	50.0	29	2	AAAI9290		AAAI9290	Integrin
20	16	50.0	34	10	ADD28468		ADD28468	Escherich
C 21	16	50.0	34	10	ADD28469		ADD28469	Escherich

FT	/note= "Forms a double-stranded region with nucleotides
FT	-2 of sequence AAL40447"
XX	
XX	JP2002119283-A.
XX	
XX	23-APR-2002.
PD	
XX	
XX	13-OCT-2000; 2000JP-00313320.
PF	
XX	
XX	13-OCT-2000; 2000JP-00313320.
PR	
XX	(DOKU-) DOKURITSU GYOSEI HOJIN SANGYO GIJUTSU SO.
XX	
XX	WPI; 2002-483792/52.
DR	
XX	
XX	A nucleic acid enzyme which has selective and effective eradicating
PT	activity towards harmful cells by acquiring cleavage activity of a
PT	specific target RNA by recognition of the other RNA molecule.
PT	
XX	Claim 2; Page 2; 17pp; Japanese.
PS	
PS	
XX	
XX	The invention relates to a nucleic acid enzyme with modifiable RNA
CC	cleavage activity. More specifically the invention relates to a nucleic
CC	acid enzyme, trans maxizyme, which has selective and effective
CC	eradicating activity towards harmful cells by acquiring cleavage activity
CC	of a specific target RNA by recognition of the other RNA molecule. The
CC	enzyme of the invention is useful for cleaving target RNA and is useful
CC	in treating diseases caused by the target RNA. This polynucleotide
CC	sequence represents the maxizyme-constituting RNA molecule T-MzR relating
CC	to the invention
XX	
XX	
SQ	Sequence 32 BP; 5 A; 7 C; 11 G; 0 T; 9 U; 0 Other;
	Query Match 100.0%; Score 32; DB 6; Length 32;
	Best Local Similarity 100.0%; Pred. No. 0.00034;
	Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 GGUCCUGGCCUGAUGAGAGUGAUGAGGCUUCUC 32
Db	1 GGUCCUGGCCUGAUGAGAGUGAUGAGGCUUCUC 32
RESULT 2	
AAL40448	
ID	AAL40448 standard; RNA; 33 BP.
XX	
AC	AAL40448;
XX	
DT	19-SBP-2002 (first entry)
XX	
XX	Maxizyme-constituting RNA molecule B-MzR.
XX	
XX	Enzyme; modifiable RNA cleavage activity; maxizyme-constituting RNA;
KW	trans maxizyme; B-MzR; ss.
XX	
OS	Unidentified.
XX	
XX	
FH	Key Location/Qualifiers
FT	misc_binding 1..9
FT	/*tag= a
FT	/bound moiety= "Bcl-2 mRNA"
FT	/note= "Forms a double-stranded region with nucleotides
FT	19-11 of sequence AAL40451"
FT	12..19
FT	misc_binding
FT	/*tag= d
FT	/bound moiety= "b-MzL RNA"
FT	/note= "Forms a double-stranded region with nucleotides
FT	17-10 of sequence AAL40449"
FT	14..19
FT	misc_binding
FT	/*tag= b
FT	/bound moiety= "B-MzL RNA"
FT	/note= "Forms a double-stranded region with nucleotides
FT	20-15 of sequence AAL40449"
FT	
FT	

```

FT misc_binding 20..33
FT /tag= c
FT /bound moiety= "BCR-ABL mRNA"
FT /note= "Forms a double-stranded region with nucleotides
FT 14-1 of sequence AAL40453"
XX
XX JP2002119283-A.
XX
XX 23-APR-2002.
XX
XX 13-OCT-2000; 2000JP-003113320.
XX
XX 13-OCT-2000; 2000JP-003113320.
XX
XX (DOKU-) DOKURITSU GYOSEI HOJIN SANGYO GIJUTSU SO.
XX
XX WPI; 2002-483792/52.
XX
XX A nucleic acid enzyme which has selective and effective eradicating
XX activity towards harmful cells by acquiring cleavage activity of a
XX specific target RNA by recognition of the other RNA molecule.
XX
XX Claim 2; Page 2; 17pp; Japanese.
XX
XX The invention relates to a nucleic acid enzyme with modifiable RNA
XX cleavage activity. More specifically the invention relates to a nucleic
XX acid enzyme, trans maxizyme, which has selective and effective
XX eradicating activity towards harmful cells by acquiring cleavage activity
XX of a specific target RNA by recognition of the other RNA molecule. The
XX enzyme of the invention is useful for cleaving target RNA and is useful
XX in treating diseases caused by the target RNA. This polynucleotide
XX sequence represents the maxizyme-constituting RNA molecule B-MzR relating
XX to the invention
XX
XX SQ Sequence 33 BP; 6 A; 5 C; 12 G; 0 T; 10 U; 0 Other;
XX
XX Query Match 66.9%; Score 21.4; DB 6; Length 33;
XX Best Local Similarity 95.7%; Pred. No. 14;
XX Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0
XX
XX QY 1 GGUCUCCGCGCUGAGAGAGUGAU 23
XX ||||||||||||||||||||
XX DB 1 GGUCUCCGCGCUGAGAGAGUGAU 23
XX
XX RESULT 3
XX ABN44085
XX ID ABN44085 standard; DNA; 60 BP.
XX
XX AC ABN44085;
XX
XX DT 15-JUL-2002 (first entry)
XX
XX DE Human spliced transcript detection oligonucleotide SEQ ID NO:16833.
XX
XX KW Human; mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200210449-A2.
XX
XX PD 07-FEB-2002.
XX
XX PF 20-JUL-2001; 2001WO-IB001903.
XX
XX PR 28-JUL-2000; 2000US-0221607P.
XX
XX PR 02-MAY-2001; 2001US-0287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;

```

```
; FastDB parameter file
; Created By tport on Mon 1 Nov 104 16:34:33-PST.

Query-File-Name      US09111292.seq
Query-Sequence-Name-or-Number  US-09-111-292-6
Query-Sequence-Range 1 1099
File-to-Search        US09111292.seq
Search-Name           us-09-111-292-6
Similarity-matrix      Unitary
Number-of-PAMs         150
Threshold-level-of-similarity 100
Ktuple                 4
Mismatch-penalty       1
Joining-penalty        30
Cutoff-score           1
Randomization-group-length 0
Number-of-randomizations 0
Translation-frame      0
Window-size            500
Gap-penalty            5.00
Gap-size-penalty       0.33
Time-to-submit-batch-job none
Number-of-initial-scores-to-save 2
Number-of-optimized-scores-to-save 20
Number-of-alignments-to-save 2
Display-annotations    No
Display-context        50
Notify-upon-completion Yes
Number-of-sequences-to-collect 13
Number-to-save-in-indirect-file 13
```



```

QY 61 UUCGAAACCGGGCAGCUACAAACCAAC 88
Db 61 TTCGAAACCGGGCAGCUACAAACCAAC 88

RESULT 12
AXI38489
LOCUS AXI38489 142 bp RNA linear PAT 30-MAY-2001
DEFINITION Sequence 50 from Patent EP1097993.
ACCESSION AXI38489
VERSION AXI38489.1 GI:14274385
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Taira,K., Warashina,M., Kuwabara,T. and Kawasaki,H.
TITLE Functional ribozyme chimeric molecules capable of sliding
JOURNAL Patent: EP 1097993-A 50 09-MAY-2001;
SECRETARY OF AGENCY OF INDUSTRIAL SCIENCE AND TECHNOLOGY (JP) ;
Taira, Kazunari (JP)
FEATURES
source
1..142
/organism="synthetic construct"
/mol_type="unassigned RNA"
/db_xref="taxon:32630"
/note="the nucleotide sequence of CPP Rz2"

ORIGIN
Query Match 100.0%; Score 88; DB 6; Length 142;
Best Local Similarity 78.4%; Pred. No. 1.4e-19;
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCGUUGUUUCGUGUAGUGGUUUAUCACGUCGCUACACGCGAAAGGUCGCCGG 60
Db 1 ACCGTGGTTTCGTAGTAGTGGTTATACGCTTCGCTAACACGCGAAAGGTCGCCGG 60

QY 61 UUCGAAACCGGGCAGCUACAAACCAAC 88
Db 61 TTCGAAACCGGGCAGCUACAAACCAAC 88

RESULT 13
AXI38490
LOCUS AXI38490 142 bp RNA linear PAT 30-MAY-2001
DEFINITION Sequence 51 from Patent EP1097993.
ACCESSION AXI38490
VERSION AXI38490.1 GI:14274386
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Taira,K., Warashina,M., Kuwabara,T. and Kawasaki,H.
TITLE Functional ribozyme chimeric molecules capable of sliding
JOURNAL Patent: EP 1097993-A 51 09-MAY-2001;
SECRETARY OF AGENCY OF INDUSTRIAL SCIENCE AND TECHNOLOGY (JP) ;
Taira, Kazunari (JP)
FEATURES
source
1..142
/organism="synthetic construct"
/mol_type="unassigned RNA"
/db_xref="taxon:32630"
/note="the nucleotide sequence of CPP Rz3"

ORIGIN
Query Match 100.0%; Score 88; DB 6; Length 142;
Best Local Similarity 78.4%; Pred. No. 1.4e-19;
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCGUUGUUUCGUGUAGUGGUUUAUCACGUCGCUACACGCGAAAGGUCGCCGG 60
Db 1 ACCGTGGTTTCGTAGTAGTGGTTATACGCTTCGCTAACACGCGAAAGGTCGCCGG 60

QY 61 UUCGAAACCGGGCAGCUACAAACCAAC 88
Db 61 TTCGAAACCGGGCAGCUACAAACCAAC 88

RESULT 14
AXI38492
LOCUS AXI38492 142 bp RNA linear PAT 30-MAY-2001
DEFINITION Sequence 53 from Patent EP1097993.
ACCESSION AXI38492
VERSION AXI38492.1 GI:14274388
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Taira,K., Warashina,M., Kuwabara,T. and Kawasaki,H.
TITLE Functional ribozyme chimeric molecules capable of sliding
JOURNAL Patent: EP 1097993-A 53 09-MAY-2001;
SECRETARY OF AGENCY OF INDUSTRIAL SCIENCE AND TECHNOLOGY (JP) ;
Taira, Kazunari (JP)
FEATURES
source
1..142
/organism="synthetic construct"
/mol_type="unassigned RNA"
/db_xref="taxon:32630"
/note="the nucleotide sequence of CPP Rz5"

ORIGIN
Query Match 100.0%; Score 88; DB 6; Length 142;
Best Local Similarity 78.4%; Pred. No. 1.4e-19;
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCGUUGUUUCGUGUAGUGGUUUAUCACGUCGCUACACGCGAAAGGUCGCCGG 60
Db 1 ACCGTGGTTTCGTAGTAGTGGTTATACGCTTCGCTAACACGCGAAAGGTCGCCGG 60

QY 61 UUCGAAACCGGGCAGCUACAAACCAAC 88
Db 61 TTCGAAACCGGGCAGCUACAAACCAAC 88

RESULT 15
BD015656
LOCUS BD015656 142 bp RNA linear PAT 27-AUG-2002
DEFINITION Slidable functional chimeric molecule.
ACCESSION BD015656
VERSION BD015656.1 GI:22556793
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 142)
AUTHORS Taira,K., Warashina,M., Kuwabara,T. and Kawasaki,H.
TITLE Slidable functional chimeric molecule
JOURNAL Patent: JP 2001190282-A 49 17-JUL-2001;
DIRECTOR GENERAL OF NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL
SCIENCE AND MANABU NAKAJIMA TECHNOLOGY, KAZUNARI TAIRA
COMMENT
OS Artificial Sequence
PN JP 2001190282-A/49
PD 17-JUL-2001
PF 02-NOV-2000 JP 2000336082
PI KAZUNARI TAIRA,MASAKI WARASHINA,TOMOKO KUWABARA,HIROAKI PI
KAWASAKI
PC
C12N15/09,A61K31/7105,A61K31/711,A61K38/00,A61K48/00,A61P31/12, PC
A61P43/00,
PC C12N9/22,C12Q1/02,C12Q1/68,G01N33/53,G01N33/566,C12N15/00, PC
A61K37/02
CC Description of Artificial Sequence: the
nucleotide sequence of

```

Mon Nov 1 16:24:21 2004

```
CC          CPP Rzl
FH Key      Location/Qualifiers
FT source   1..142
            /organism='Artificial Sequence'
FEATURES
  source    Location/Qualifiers
            1..142
            /organism="synthetic construct"
            /mol_type="genomic RNA"
            /db_xref="taxon:32630"
```

ORIGIN

```
Query Match      100.0%; Score 88; DB 6; Length 142;
Best Local Similarity 78.4%; Pred. No. 1.4e-19;
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCGUUGGUUCCGUAGUGUGUAUCACGUUCCCUAACACGCGAAGGUCGCCGG 60
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 1 ACCGTTGGTTTCCGCGTAGTGTATCAGTTTCGCTAACACGCGAAGGTCGCCGG 60
   |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

Qy 61 UUCGAAACCGGGCACUACAAAACCAAC 88
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db 61 TTCGAAACCGGGCGACTACAAAACCAAC 88
   :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
```

Search completed: October 31, 2004, 04:50:41  
Job time : 1507 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 31, 2004, 03:24:30 ; Search time 1505 Seconds  
(without alignments)  
2765.112 Million cell updates/sec

Title: US-09-974-974-5

Perfect score: 88

Sequence: 1 accguugnuccguagu.....cgggcacacaaacaaac 88

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 segs, 23644849745 residues

Total number of hits satisfying chosen parameters: 2719732

Minimum DB seq length: 0

Maximum DB seq length: 200

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_scs.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	88	100.0	88	6	BD182356	BD182356 Novel max
2	88	100.0	88	6	AX453846	AX453846 Sequence
3	88	100.0	88	6	BD143502	BD143502 Nucleic a
4	88	100.0	100	6	E47174	E47174 Method for
5	88	100.0	117	6	BD174675	BD174675 Ribozyme
6	88	100.0	137	6	AX429079	AX429079 Sequence
7	88	100.0	137	6	BD143601	BD143601 Method of
8	88	100.0	138	6	AX453858	AX453858 Sequence
9	88	100.0	141	6	AX138491	AX138491 Sequence
10	88	100.0	141	6	BD015659	BD015659 Slidable
11	88	100.0	142	6	AX138488	AX138488 Sequence
12	88	100.0	142	6	AX138489	AX138489 Sequence
13	88	100.0	142	6	AX138490	AX138490 Sequence
14	88	100.0	142	6	AX138492	AX138492 Sequence
15	88	100.0	142	6	BD015656	BD015656 Slidable
16	88	100.0	142	6	BD015657	BD015657 Slidable
17	88	100.0	142	6	BD015658	BD015658 Slidable
18	88	100.0	142	6	BD015660	BD015660 Slidable
19	88	100.0	151	6	AX138449	AX138449 Sequence

20	88	100.0	151	6	BD015617	BD015617 Slidable
21	88	100.0	175	6	BD174679	BD174679 Ribozyme
22	88	100.0	175	6	BD174680	BD174680 Ribozyme
23	82.8	94.1	136	6	E33203	E33203 Expression
24	82.8	94.1	153	12	AB080619	AB080619 Synthetic
25	82.8	94.1	153	12	AB080620	AB080620 Synthetic
26	82.8	94.1	153	12	AB080621	AB080621 Synthetic
27	82.8	94.1	153	12	AB080622	AB080622 Synthetic
28	82.8	94.1	153	12	AB080623	AB080623 Synthetic
29	82.8	94.1	153	12	AB080624	AB080624 Synthetic
30	81	92.0	142	6	E33204	E33204 Expression
31	76	86.4	128	6	E33205	E33205 Expression
32	74.8	85.0	95	6	E33206	E33206 Expression
33	74.8	85.0	95	6	AX138447	AX138447 Sequence
34	74.8	85.0	95	6	BD015615	BD015615 Slidable
c						
35	74	84.1	183	11	G31283	G31283 sy905g1-L H
36	73.4	83.4	149	6	E33207	E33207 Expression
37	70.4	80.0	132	6	AR041184	AR041184 Sequence
38	70.4	80.0	132	6	AR113038	AR113038 Sequence
39	70.4	80.0	132	6	I66463	I66463 Sequence 17
c						
40	66.8	75.9	113	6	E33209	E33209 Expression
41	66	75.0	110	6	E33208	E33208 Expression
42	65	73.9	75	10	RATRV1	K00256 Rat liver V
43	64.4	73.2	76	4	RATRV1	K00253 Rabbit Val-
44	64.4	73.2	76	9	HUMTRV1A	K00255 Human Val-t
45	64.4	73.2	76	9	HUMTRV1B	K00254 Human Val-t

## ALIGNMENTS

RESULT 1	BD182356	88 bp	RNA	linear	PAT 15-MAY-2003
LOCUS	Novel maxizyme.				
DEFINITION	BD182356				
ACCESSION	BD182356				
VERSION	BD182356.1 GI:30793274				
KEYWORDS	WO 02092821-A/3.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 88)

Taira, K., Warashina, T., Warashina, M., Kawasaki, H., Hara, T. and Nozawa, I.

Novel maxizyme

Patent: WO 02092821-A 3 21-NOV-2002;

NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY,

TAISHO PHARMACEUTICAL CO LTD, GENOFUNCTION INC, KAZUNARI TAIRA,

TOMOKO WARASHINA, MASAKI WARASHINA, HIROAKI KAWASAKI, TOSHIYUKI HARA,

IWAO NOZAWA

OS Homo sapiens (human)

PN WO 02092821-A/3

PD 21-NOV-2002

PF 30-APR-2002 WO 2002JP004322

PR 01-MAY-2001 JP 01P 134469

PI KAZUNARI TAIRA, TOMOKO WARASHINA, MASAKI WARASHINA, HIROAKI PI

KAWASAKI,

PI TOSHIYUKI HARA, IWAO NOZAWA

PC C12N15/55, C12N9/22, C12Q1/34, A61K48/00, A61K38/43 CC Novel

maxizyme

PH Key

FT source

Location/Qualifiers

1. .87

/organism='Homo sapiens (human)'

1. 88

/organism='Homo sapiens'

/mol\_type='genomic RNA'

/db\_xref='taxon:9606'

ORIGIN

Query Match 100.0%; Score 88; DB 6; Length 88;

Best Local Similarity 78.4%; Pred. No. 1.4e-19;

C12N15/09,A61K9/127,A61K38/46,A61K48/00,A61P31/12,A61P35/00, PC  
C12N1/15,  
PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12Q1/25,C12Q1/68,C12N15/1/  
00,A61K37/54,  
PC C12N5/00  
CC Description of Artificial Sequence: tRNAval promoter sequence  
FH Key Location/Qualifiers  
FT source 1..88  
/organism='Artificial Sequence'.  
FT Location/Qualifiers  
1..88  
/organism='synthetic construct'  
/mol\_type='genomic RNA'  
/db\_xref='taxon:32630'

# FEATURES

source

## ORIGIN

Query Match 100.0%; Score 88; DB 6; Length 88;  
Best Local Similarity 78.4%; Pred. No. 1.4e-19;  
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCGUUGGUUCCGUAGUGUAGUGUUAUCACGUCGCUAACACGCGAAAGGUCCCGG 60  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 ACCGTTGGTTCCGTAGTGTATCAGTTGCTTACACGCGAAAGGTCCCGG 60  
QY 61 UUCGAAACCGGGCACUACAAAACCAAC 88  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 TTCGAAACCGGGCACTACAAAACCAAC 88

## RESULT 4

E47174 100 bp DNA linear PAT 27-AUG-2002  
LOCUS Method for selecting active ribozyme.  
DEFINITION  
ACCESSION E47174  
VERSION E47174.1 GI:22553362  
KEYWORDS JP 2001128682-A/9.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 100)  
Tabira,K. and Kawasaki,H.  
Method for selecting active ribozyme  
Patent: JP 2001128682-A 9 15-MAY-2001;  
JOURNAL AGENCY OF IND SCIENCE & TECHNOL

## COMMENT

OS Homo sapiens (human)  
PN JP 2001128682-A/9  
PD 15-MAY-2001  
PF 05-NOV-1999 JP 1999314579  
PI KAZUNARI TABIRA,HIROAKI KAWASAKI  
PC C12N15/09,C12Q1/68//C12N5/10,C12N15/00,C12N5/00 CC  
FH Key Location/Qualifiers.

## FEATURES

source  
1..100  
/organism='Homo sapiens'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

## ORIGIN

Query Match 100.0%; Score 88; DB 6; Length 100;  
Best Local Similarity 78.4%; Pred. No. 1.4e-19;  
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCGUUGGUUCCGUAGUGUAGUGUUAUCACGUCGCUAACACGCGAAAGGUCCCGG 60  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 ACCGTTGGTTCCGTAGTGTATCAGTTGCTTACACGCGAAAGGTCCCGG 60  
QY 61 UUCGAAACCGGGCACUACAAAACCAAC 88  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 TTCGAAACCGGGCACTACAAAACCAAC 88

## RESULT 5

Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ACCGUUGGUUCCGUAGUGUAGUGUUAUCACGUCGCUAACACGCGAAAGGUCCCGG 60  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 ACCGTTGGTTCCGTAGTGTATCAGTTGCTTACACGCGAAAGGTCCCGG 60  
QY 61 UUCGAAACCGGGCACUACAAAACCAAC 88  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 TTCGAAACCGGGCACTACAAAACCAAC 88

## RESULT 2

AX453846 88 bp RNA linear PAT 06-JUL-2002  
LOCUS Sequence 5 from Patent EP1213351.  
DEFINITION  
ACCESSION AX453846  
VERSION AX453846.1 GI:21713515  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1  
AUTHORS Taira,K., Warashina,M. and Warashina,T.  
TITLE Nucleic acid enzymes acquiring an activity for cleaving a target  
rna by recognising another molecule  
JOURNAL Patent: EP 1213351-A 5 12-JUN-2002;  
National Institute of Advanced Industrial Science and Technology  
(JP)

## FEATURES

source  
1..88  
/organism='synthetic construct'  
/mol\_type='unassigned RNA'  
/db\_xref='taxon:32630'  
/note='tRNAval promoter sequence'

## ORIGIN

Query Match 100.0%; Score 88; DB 6; Length 88;  
Best Local Similarity 78.4%; Pred. No. 1.4e-19;  
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCGUUGGUUCCGUAGUGUAGUGUUAUCACGUCGCUAACACGCGAAAGGUCCCGG 60  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 ACCGTTGGTTCCGTAGTGTATCAGTTGCTTACACGCGAAAGGTCCCGG 60  
QY 61 UUCGAAACCGGGCACUACAAAACCAAC 88  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 TTCGAAACCGGGCACTACAAAACCAAC 88

## RESULT 3

BD143502 88 bp RNA linear PAT 17-JAN-2003  
LOCUS Nucleic acid enzyme acquiring activity of cleaving other specific  
DEFINITION target RNA by recognizing RNA molecule.  
ACCESSION BD143502  
VERSION BD143502.1 GI:27849260  
KEYWORDS JP 2002119283-A/5.  
SOURCE synthetic construct  
ORGANISM artificial sequences.  
REFERENCE 1 (bases 1 to 88)  
AUTHORS Taira,K., Warashina,M. and Warashina,T.  
TITLE Nucleic acid enzyme acquiring activity of cleaving other specific  
JOURNAL target RNA by recognizing RNA molecule  
Patent: JP 2002119283-A 5 23-APR-2002;  
DIRECTOR GENERAL OF NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL  
SCIENCE AND P T BAKRIE AND BROTHERS,JAPAN BIO INDUSTRY ASSOCIATION,  
TECHNOLOGY  
OS Artificial Sequence  
PN JP 2002119283-A/5  
PD 23-APR-2002  
PF 13-OCT-2000 JP 2000313320  
PI KAZUNARI TAIRA,MASAKI WARASHINA,TOMOKO WARASHINA PC

```
BD174675
LOCUS BD174675 117 bp DNA linear PAT 18-MAR-2003
DEFINITION Ribozyme expression system.
ACCESSION BD174675
VERSION BD174675.1 GI:29120365
KEYWORDS JP 2002262880-A/1.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 117)
AUTHORS Takebe, Y. and Okawa, J.
TITLES Ribozyme expression system
JOURNAL Patent: JP 2002262880-A 1 17-SEP-2002;
DIRECTOR GENERAL OF NATIONAL INSTITUTE OF INFECTIONS DISEASES,
YUTAKA TAKEBE, JUN OKAWA
OS Artificial Sequence
PN JP 2002262880-A/1
PD 17-SEP-2002
PF 09-MAR-2001 JP 2001067253
PI YUTAKA TAKEBE, JUN OKAWA
PC C12N15/09, A61K31/711, A61K35/76, A61K48/00, A61P31/18, C12N5/10,
PC C12N9/00,
PC C12N15/00, C12N5/00
CC Description of Artificial Sequence: the nucleotide sequence of
CC tetO1-trNAval
FH Key Location/Qualifiers
FT source 1. .117
FEATURES
source Location/Qualifiers
1. .117
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 100.0%; Score 88; DB 6; Length 117;
Best Local Similarity 78.4%; Pred. No. 1.4e-19;
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCGUUGUUUCCGUAGUGUAGUUCACGUUCGCCUUAACACGCGAAAGGUCCCGG 60
DB 29 ACCGTGGTTTCGTAGTGTAGTGGTTATCAGTTCGCTTAACACGCGAAAGGTCCCGG 88
QY 61 UUCGAAACCGGGCACUACAAAACCAAC 88
DB 89 TTCGAAACCGGGCACTACAAAACCAAC 116
RESULT 6
AX429079
LOCUS AX429079 137 bp RNA linear PAT 21-JUN-2002
DEFINITION Sequence 2 from Patent EP1201751.
ACCESSION AX429079
VERSION AX429079.1 GI:21540419
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Taira, K. and Sano, M.
TITLES Method for selecting highly functional nucleic acid molecules
JOURNAL Patent: EP 1201751-A 2 02-MAY-2002;
National Institute of Advanced Industrial Science and Technology
(JP)
FEATURES
source Location/Qualifiers
1. .137
/organism="synthetic construct"
/mol_type="unassigned RNA"
/db_xref="taxon:32630"
/note="Sequence of tRNA-Luc RZ"
ORIGIN
Query Match 100.0%; Score 88; DB 6; Length 137;
Best Local Similarity 78.4%; Pred. No. 1.4e-19;
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCGUUGUUUCCGUAGUGUAGUUCACGUUCGCCUUAACACGCGAAAGGUCCCGG 60
DB 29 ACCGTGGTTTCGTAGTGTAGTGGTTATCAGTTCGCTTAACACGCGAAAGGTCCCGG 88
QY 61 UUCGAAACCGGGCACUACAAAACCAAC 88
DB 89 TTCGAAACCGGGCACTACAAAACCAAC 116
RESULT 7
AX453858
LOCUS AX453858 138 bp RNA linear PAT 06-JUL-2002
DEFINITION Sequence 17 from Patent EP1213351.
ACCESSION AX453858
VERSION AX453858.1 GI:21713527
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Taira, K., Warashina, M. and Warashina, T.
TITLES Nucleic acid enzymes acquiring an activity for cleaving a target
```

```
Query Match 100.0%; Score 88; DB 6; Length 137;
Best Local Similarity 78.4%; Pred. No. 1.4e-19;
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCGUUGUUUCCGUAGUGUAGUUCACGUUCGCCUUAACACGCGAAAGGUCCCGG 60
DB 1 ACCGTGGTTTCGTAGTGTAGTGGTTATCAGTTCGCTTAACACGCGAAAGGTCCCGG 60
QY 61 UUCGAAACCGGGCACUACAAAACCAAC 88
DB 61 TTCGAAACCGGGCACTACAAAACCAAC 88
RESULT 7
BD143601
LOCUS BD143601 137 bp RNA linear PAT 17-JAN-2003
DEFINITION Method of selecting high-function nucleic acid molecule in cell.
ACCESSION BD143601
VERSION BD143601.1 GI:27849359
KEYWORDS JP 2002125685-A/2.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 137)
AUTHORS Taira, K. and Sano, M.
TITLES Method of selecting high-function nucleic acid molecule in cell
JOURNAL Patent: JP 2002125685-A 2 08-MAY-2002;
DIRECTOR GENERAL OF NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL
SCIENCE AND HIROSHI YATSUHASHI, TECHNOLOGY, KAZUNARI TAIRA
OS Artificial Sequence
PN JP 2002125685-A/2
PD 08-MAY-2002 JP 2000331347
PF 30-OCT-2000 JP 2000331347
PI KAZUNARI TAIRA, MASAYUKI SANO
PC C12N15/09, C12N9/00, C12Q1/02, C12Q1/25, C12Q1/68, C12N15/00 CC
CC Description of Artificial Sequence: Sequence of tRNA-Luc GUA CC
FH Key RZ Location/Qualifiers
FT source 1. .137
FEATURES
source Location/Qualifiers
1. .137
/organism="synthetic construct"
/mol_type="genomic RNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 100.0%; Score 88; DB 6; Length 137;
Best Local Similarity 78.4%; Pred. No. 1.4e-19;
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACCGUUGUUUCCGUAGUGUAGUUCACGUUCGCCUUAACACGCGAAAGGUCCCGG 60
DB 1 ACCGTGGTTTCGTAGTGTAGTGGTTATCAGTTCGCTTAACACGCGAAAGGTCCCGG 60
QY 61 UUCGAAACCGGGCACUACAAAACCAAC 88
DB 61 TTCGAAACCGGGCACTACAAAACCAAC 88
RESULT 8
AX453858
LOCUS AX453858 138 bp RNA linear PAT 06-JUL-2002
DEFINITION Sequence 17 from Patent EP1213351.
ACCESSION AX453858
VERSION AX453858.1 GI:21713527
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Taira, K., Warashina, M. and Warashina, T.
TITLES Nucleic acid enzymes acquiring an activity for cleaving a target
```

rna by recognising another molecule  
Patent: EP 1213351-A 17 12-JUN-2002;  
National Institute of Advanced Industrial Science and Technology  
(JP)

FEATURES  
source  
1. .138  
/organism="synthetic construct"  
/mol\_type="unassigned RNA"  
/db\_xref="taxon:32630"  
/note="tRNAVal T-MzL"

ORIGIN  
Query Match 100.0%; Score 88; DB 6; Length 138;  
Best Local Similarity 78.4%; Pred. No. 1.4e-19;  
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCGUUGGUUCCGUAGUGUUAUCACGUCGUUUAUCACGCGGAAAGGUCCCGG 60  
Db 1 ACCGTTGGTTCCGTTAGTGTAGTTATCAGTTGCTTAACACGCGGAAAGGTCCCGG 60

Qy 61 UUCGAAACCGGGCACUACAAAACCAAC 88  
Db 61 TTCGAAACCGGGCACTACAAAACCAAC 88

RESULT 9  
AXI38491 141 bp RNA linear PAT 30-MAY-2001  
LOCUS  
DEFINITION Sequence 52 from Patent EPI097993.  
AXI38491  
ACCESSION AXI38491  
VERSION AXI38491.1 GI:14274387  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.

REFERENCE  
1  
AUTHORS Taira, K., Warashina, M., Kuwabara, T. and Kawasaki, H.  
TITLE Functional ribozyme chimeric molecules capable of sliding  
JOURNAL Patent: EP 1097993-A 52 09-MAY-2001;  
Secretary of Agency of Industrial Science and Technology (JP);  
Taira, Kazunari (JP)

FEATURES  
source  
1. .141  
Location/Qualifiers  
/organism="synthetic construct"  
/mol\_type="unassigned RNA"  
/db\_xref="taxon:32630"  
/note="the nucleotide sequence of CPP Rz4"

ORIGIN  
Query Match 100.0%; Score 88; DB 6; Length 141;  
Best Local Similarity 78.4%; Pred. No. 1.4e-19;  
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCGUUGGUUCCGUAGUGUUAUCACGUCGUUUAUCACGCGGAAAGGUCCCGG 60  
Db 1 ACCGTTGGTTCCGTTAGTGTAGTTATCAGTTGCTTAACACGCGGAAAGGTCCCGG 60

Qy 61 UUCGAAACCGGGCACUACAAAACCAAC 88  
Db 61 TTCGAAACCGGGCACTACAAAACCAAC 88

RESULT 10  
BD015659 141 bp RNA linear PAT 27-AUG-2002  
LOCUS  
DEFINITION Slidable functional chimeric molecule.  
BD015659  
ACCESSION BD015659  
VERSION BD015659.1 GI:22556796  
KEYWORDS JP 2001190282-A/52.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.  
1 (bases 1 to 141)

AUTHORS  
TITLE Slidable functional chimeric molecule  
JOURNAL Patent: EP 2001190282-A 52 17-JUL-2001;  
DIRECTOR GENERAL OF NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL  
SCIENCE AND MANABU NAKAJIMA TECHNOLOGY, KAZUNARI TAIRA  
OS Artificial Sequence  
PN JP 2001190282-A/52  
PD 17-JUL-2001  
PF 02-NOV-2000 JP 2000336082  
PI KAZUNARI TAIRA, MASAKI WARASHINA, TOMOKO KUWABARA, HIROAKI  
KAWASAKI  
PC C12N15/09, A61K31/7105, A61K31/711, A61K38/00, A61K48/00, A61P31/12, PC  
A61P43/00,  
PC C12N9/22, C12Q1/02, C12Q1/68, G01N33/53, G01N33/566, C12N15/00, PC  
A61K37/02  
CC Description of Artificial Sequence: the  
nucleotide sequence of  
CPP Rz4  
FH Key Location/Qualifiers  
FT source 1. .141  
/organism="Artificial Sequence".  
FEATURES  
source  
1. .141  
Location/Qualifiers  
/organism="synthetic construct"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:32630"

ORIGIN  
Query Match 100.0%; Score 88; DB 6; Length 141;  
Best Local Similarity 78.4%; Pred. No. 1.4e-19;  
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCGUUGGUUCCGUAGUGUUAUCACGUCGUUUAUCACGCGGAAAGGUCCCGG 60  
Db 1 ACCGTTGGTTCCGTTAGTGTAGTTATCAGTTGCTTAACACGCGGAAAGGTCCCGG 60

Qy 61 UUCGAAACCGGGCACUACAAAACCAAC 88  
Db 61 TTCGAAACCGGGCACTACAAAACCAAC 88

RESULT 11  
AXI38488 142 bp RNA linear PAT 30-MAY-2001  
LOCUS  
DEFINITION Sequence 49 from Patent EPI097993.  
AXI38488  
ACCESSION AXI38488  
VERSION AXI38488.1 GI:14274384  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
artificial sequences.

REFERENCE  
1  
AUTHORS Taira, K., Warashina, M., Kuwabara, T. and Kawasaki, H.  
TITLE Functional ribozyme chimeric molecules capable of sliding  
JOURNAL Patent: EP 1097993-A 49 09-MAY-2001;  
Secretary of Agency of Industrial Science and Technology (JP);  
Taira, Kazunari (JP)

FEATURES  
source  
1. .142  
Location/Qualifiers  
/organism="synthetic construct"  
/mol\_type="unassigned RNA"  
/db\_xref="taxon:32630"  
/note="the nucleotide sequence of CPP Rz1"

ORIGIN  
Query Match 100.0%; Score 88; DB 6; Length 142;  
Best Local Similarity 78.4%; Pred. No. 1.4e-19;  
Matches 69; Conservative 19; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACCGUUGGUUCCGUAGUGUUAUCACGUCGUUUAUCACGCGGAAAGGUCCCGG 60  
Db 1 ACCGTTGGTTCCGTTAGTGTAGTTATCAGTTGCTTAACACGCGGAAAGGTCCCGG 60





BioTechniques 35 (6), 1164-1168 (2003)  
14682050  
4 (bases 1 to 43)  
Li.Y., Rosso,M.G., Strizhov,N. and Weishaar,B.  
Direct Submission  
Submitted (31-MAR-2004) Weisshaar B., Max-Planck-Institut fuer  
Zuchtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany  
This sequence has been recovered from the left border of the T-DNA.  
It indicates an insertion within the locus defined by BAC clone  
f25p17. Details on the protocols used for generation of the  
sequence are described in References 1-3. The sequences are  
generated at the MPI for Plant Breeding Research in the context of  
the GABI-Kat project. GABI-Kat is part of the German Plant Genomics  
program designated 'GABI'. Information on line availability can be  
found at: http://www.mpiz-koeln.mpg.de/GABI-Kat/  
Location/Qualifiers

**FEATURES**

source	1..43
	/organism="Arabidopsis thaliana"
	/mol_type="genomic DNA"
	/strain="Columbia 0"
	/db_xref="taxon:3702"
	/clone="GK-144A04-012978"
	/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
	/ecotype="Col-0"
	/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161 (GenBank accession number: AJ537514). The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed."

**ORIGIN**

Query Match	51.1%;	Score 13.8;	DB 9;	Length 43;
Best Local Similarity	44.0%;	Pred. No. 7.9e+04;		
Matches	11; Conservative	7; Mismatches	7; Indels	0; Gaps 0;

**QY**

Db	2 UCUGACUGUUAUCGAACCGGGC 26
	:   :::   :
	6 TCTGTCTGTGTTTAAAGTGCGTC 30

**RESULT 5**

AV834227/c

**LOCUS**

AV834227 K. Sato unpublished cDNA library: Hordeum vulgare subsp. vulgare shoots germination Hordeum vulgare subsp. vulgare cdna clone bags/j23, mRNA sequence.

**DEFINITION**

AV834227

**ACCESSION**

AV834227.1 GI:14526316

**VERSION**

EST.

**KEYWORDS**

SOURCE

ORGANISM

Hordeum vulgare subsp. vulgare  
Hordeum vulgare subsp. vulgare  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Pooidae; Triticeae; Hordeum.  
1 (bases 1 to 48)  
Sato,K.  
Barley EST sequencing project in NIG and Okayama Univ  
Unpublished (2001)  
Contact: Kazuhiro Sato  
Research Institute for Bioresources  
Okayama University, Barley Germplasm Center  
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan  
Email: karsato@rib.okayama-u.ac.jp,  
URL:http://www.rib.okayama-u.ac.jp/barley/  
database:http://www.shigen.nig.ac.jp/barley.html.  
Location/Qualifiers

**REFERENCE**

AUTHORS	1..48
TITLE	/organism="Hordeum vulgare subsp. vulgare"
JOURNAL	/mol_type="mRNA"
COMMENT	/cultivar="Haruna Nijo"
	/sub_species="vulgare"

**FEATURES**

source





```

02S0288-04, Primer set: B
Class: transposon insertion site.
Location/Qualifiers
1. .43
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="W22 (ACR, bz1-m9)"
/cultivar="UniformMu"
/db_xref="taxon:4577"
/clone="02S0288-04B1-A02"
/clone_lib="UniformMu MUTAIL Library"
/note="Vector: TOPO-PCR4; DNA flanking Mu transposon
insertions in Mu inactive lines were extracted from the
UniformMu maize population by the thermo asymmetric
interlaced PCR (TAIL) protocol using primers specific for
the Mu terminal inverted repeat and a set of 16 arbitrary
primers. Amplicons were size enriched using Sepharose 400
spin columns and cloned into the TOPO PCR4 vector."

ORIGIN
Query Match 48.9%; Score 13.2; DB 9; Length 43;
Best Local Similarity 50.0%; Pred. No. 1.6e+05;
Matches 13; Conservative 5; Mismatches 8; Indels 0; Gaps 0.

QY 1 GUCGACUCGUCAUCCGAAACCGGUC 26
Db 3 GCCTCCATTTCGTGCAATCCGGTTC 28

RESULT 11
A2820793/c
LOCUS
DEFINITION
45 bp DNA linear GSS 20-FEB-2001
clone UUGC2M0093B18 F, genomic survey sequence.
ACCESSION
A2820793
VERSION
A2820793.1 GI:12990701
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 45)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamal,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0093 row: B column: 18
Seq primer: GGTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 45.
Location/Qualifiers
1. .45
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0093B18"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.

FEATURES
source

```



```

RESULT 14
AA905188/c
LOCUS
DEFINITION
  AA905188
  43 bp mRNA linear EST 09-JUN-1998
  IMAGE:1507373 3' similar to SW:PDK_HUMAN O00764 PYRIDOXINE KINASE
  ;, mRNA sequence.
ACCESSION
  AA905188
  GI:3040311
VERSION
  AA905188
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 43)
AUTHORS
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
JOURNAL
  Contact: Robert Strausberg, Ph.D.
  Email: cgaps-r@mail.nih.gov
COMMENT
  This clone is available royalty-free through LMLL ; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Trace considered overall poor quality
  Insert Length: 530 Std Error: 0.00
  Seq primer: -40m13 fwd. Bt from Amersham
  High quality sequence stop: 1.
FEATURES
  source
  1..43
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="IMAGE:1507373"
  /lab_host="DH10B"
  /clone_lib="Soares NFL T_GBC S1"
  /note="Organ: pooled; Vector: pT73B-Pac (Pharmacia) with
  a modified polylinker; Site 1: Not I; Site 2: Eco RI;
  Equal amounts of plasmid DNA from three normalized
  libraries (fetal lung NBHL19W, testis NHT, and B-cell
  NCI CGAP GC31) were mixed, and ss circles were made in
  vitro. Following HAP purification, this DNA was used as
  tracer in a subtractive hybridization reaction. The driver
  was PCR-amplified cDNAs from pools of 5,000 clones made
  from the same 3 libraries. The pools consisted of
  I.M.A.G.E. clones 297480-302087, 682632-687239,
  726408-728711, and 729096-731399. Subtraction by Bento
  Soares and M. Fatima Bonaldo."
ORIGIN
  Query Match 48.1%; Score 13; DB 1; Length 43;
  Best Local Similarity 47.6%; Pred. No. 1.9e+05;
  Matches 10; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 GUCUGACUGUUCAGCGAACCC 21
  ||:|:|:|:|:|:|
Db 34 GTCCTCCCTGTTGATGGAACCC 14

RESULT 15
BZ594864
LOCUS
DEFINITION
  BZ594864
  44 bp DNA linear GSS 07-JAN-2003
  SALK_085323.43.20.x Arabidopsis thaliana TDNA insertion lines
  Arabidopsis thaliana genomic clone SALK_085323.43.20.x, genomic
  survey sequence.
ACCESSION
  BZ594864
  GI:27535383
VERSION
  BZ594864.1
KEYWORDS
  GSS.
SOURCE
  Arabidopsis thaliana (thale cress)
ORGANISM
  Arabidopsis thaliana
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
  1 (bases 1 to 44)
AUTHORS
  Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
  Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
  Shinn,P., Zimmerman,J. and Ecker,J.R.
  A Sequence-Indexed Library of Insertion Mutations in the
  Arabidopsis Genome
  Unpublished (2001)
JOURNAL
  Contact: Joseph R. Ecker
  Salk Institute Genomic Analysis Laboratory (SIGAL)
  The Salk Institute for Biological Studies
  10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
  Tel: 858 453 4100 x1752
  Fax: 858 558 6379
  Email: ecker@salk.edu
COMMENT
  This is single pass sequence recovered from the left border of
  TDNA. This sequence lies within an annotated exon of At5g35690.
  Class: TDNA tagged.
FEATURES
  source
  1..44
  /organism="Arabidopsis thaliana"
  /mol_type="genomic DNA"
  /ecotype="Col-0"
  /db_xref="taxon:3702"
  /clone="SALK_085323.43.20.x"
  /clone_lib="Arabidopsis thaliana TDNA insertion lines"
  /note="PCR was performed on Arabidopsis thaliana lines
  each of which contains one or more TDNA insertion
  elements. The resultant fragment for each line was
  directly sequenced to determine the genomic sequence at
  the site of insertion. Details of the protocols used can
  be found at http://signal.salk.edu/tdna\_protocols.html"
ORIGIN
  Query Match 48.1%; Score 13; DB 8; Length 44;
  Best Local Similarity 52.4%; Pred. No. 1.9e+05;
  Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 5 GACUGUUCAGCGAACCCGGU 25
  ||:|:|:|:|:|
Db 13 GAGTGTTCATCAGAAAGGGT 33

Search completed: October 31, 2004, 04:14:10
Job time : 1373.36 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 31, 2004, 04:14:22 ; Search time 1728 Seconds  
(without alignments)  
1855.727 Million cell updates/sec

Title: US-09-974-974-5

Perfect score: 88

Sequence: 1 accggguuuccgagugu.....cgggacacacacacacac 88

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 1821985908 residues

Total number of hits satisfying chosen parameters: 3813152

Minimum DB seq length: 0

Maximum DB seq length: 200

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	40.4	45.9	57	2	BF031637	BF031637 60158122
2	39	44.3	158	8	CC059191	CC059191 i120602.b
3	38.2	43.4	171	9	AJ591533	AJ591533 Arabidops
4	33	37.5	174	8	CC004711	CC004711 PUDGC26ID
5	31.6	35.9	177	9	BX176513	BX176513 Danio rer
6	30.8	35.0	126	7	H04713	H04713 el17-1.3.3-3
7	29.8	33.9	71	9	BX203515	BX203515 Danio rer
8	23.6	33.6	150	5	BU275850	BU275850 Cr Emb.08
9	29.4	33.4	117	6	CD768454	CD768454 AGENCOURT
10	29.2	33.2	130	4	BI845249	BI845249 fs92a07.Y
11	29.2	33.2	166	8	BH330762	BH330762 CH230-141
12	29	33.0	157	8	AZ261468	AZ261468 RPCI-23-1
13	29	33.0	166	8	AZ118437	AZ118437 RPCI-23-4
14	29	33.0	175	9	BX244513	BX244513 Danio rer
15	28.8	32.7	158	7	CO739878	CO739878 SLLB06a25
16	28.8	32.7	166	7	CO739859	CO739859 SLLB06a25
17	28.4	32.3	158	4	BI708958	BI708958 fp94g05.Y
18	28.4	32.3	169	8	BH812737	BH812737 SALK 0629
19	28.4	32.3	195	4	BM532607	BM532607 fx38c08.Y
20	28.2	32.0	108	7	CO739537	CO739537 SLLB06a24
21	28.2	32.0	118	7	CO739528	CO739528 SLLB06a24
22	28.2	32.0	129	7	CO739710	CO739710 SLLB06a24
23	28.2	32.0	129	7	CO739745	CO739745 SLLB06a24
24	28.2	32.0	132	7	CO739585	CO739585 SLLB06a24

25	28.2	32.0	133	7	CO739781	CO739781 SLLB06a24
26	28.2	32.0	134	7	CO739672	CO739672 SLLB06a24
27	28.2	32.0	135	7	CO739592	CO739592 SLLB06a24
28	28.2	32.0	135	7	CO739661	CO739661 SLLB06a24
29	28.2	32.0	135	7	CO739756	CO739756 SLLB06a24
30	28.2	32.0	135	7	CO740043	CO740043 SLLB06a25
31	28.2	32.0	136	7	CO739581	CO739581 SLLB06a24
32	28.2	32.0	136	7	CO739594	CO739594 SLLB06a24
33	28.2	32.0	136	7	CO739620	CO739620 SLLB06a24
34	28.2	32.0	136	7	CO739681	CO739681 SLLB06a24
35	28.2	32.0	136	7	CO739794	CO739794 SLLB06a24
36	28.2	32.0	136	7	CO739992	CO739992 SLLB06a25
37	28.2	32.0	137	7	CO739512	CO739512 SLLB06a23
38	28.2	32.0	137	7	CO739733	CO739733 SLLB06a24
39	28.2	32.0	137	7	CO739810	CO739810 SLLB06a25
40	28.2	32.0	137	7	CO740001	CO740001 SLLB06a25
41	28.2	32.0	138	7	CO739739	CO739739 SLLB06a24
42	28.2	32.0	138	7	CO739925	CO739925 SLLB06a25
43	28.2	32.0	139	7	CO739747	CO739747 SLLB06a24
44	28.2	32.0	139	7	CO739827	CO739827 SLLB06a25
45	28.2	32.0	140	7	CO739573	CO739573 SLLB06a24

#### ALIGNMENTS

RESULT 1  
BF031637  
LOCUS  
DEFINITION  
60158122F1 NIH\_MGC\_58 Homo sapiens cDNA clone IMAGE:3827860 5',  
mRNA sequence.  
ACCESSION  
BF031637  
VERSION  
BF031637.1  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 57)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
AUTHORS  
National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE  
Unpublished (1999)  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: CLONETECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
Clone Sequencing by: Incyte Genomics, Inc.  
DNA distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLC497 row: 1 column: 05  
High quality sequence stop: 57.  
Location/Qualifiers  
1. 57  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3827860"  
/tissue\_type="hypertrophoma"  
/lab\_host="DHI0B (TL phage-resistant)"  
/clone\_lib="NIH\_MGC\_58"  
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site\_1: SfiI (ggcgctcgcc); Site\_2: SfiI (ggcattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCTAGGCGGCGGCACATG-3' and 3' adaptor (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.35 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

```

AJ591533
LOCUS
DEFINITION
Arabidopsis thaliana T-DNA flanking sequence.
587A06, genomic survey sequence.
AJ591533
ACCESSION
VERSION
GI:37941157
KEYWORDS
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1
REFERENCE
AUTHORS
Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F.,
Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
Lepiniec,L., Caboche,M. and Lecharny,A.
TITLE
T-DNA integration into the Arabidopsis genome depends on sequences
of pre-insertion sites
JOURNAL
EMBO Rep. 3 (12), 1152-1157 (2002)
MEDLINE
22363535
PUBMED
12446565
REFERENCE
2 (bases 1 to 171)
AUTHORS
Balzergue,S.
TITLE
Direct Submission
JOURNAL
Submitted (23-OCT-2003) Balzergue S., UMRGV, INRA/CNRS, 2 rue
Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT
PCR was performed on DNA from transformants of Arabidopsis thaliana
plants from INRA (Versailles). The DNA fragment(s) resulting from
the PCR were directly sequenced from the left or the right border
to determine the genomic sequence flanking the insertion. T-DNA
derived sequences were removed. Information to order the
corresponding mutant line and a link to a database providing a
graphical display of the insertion site are available at
http://dbgap.versailles.inra.fr/publiclines/. This sequence has
been generated in the framework of the French plant genomics
program 'Genoplante' (http://www.genoplante.com and
http://genoplante-info.infobiogen.fr).

FEATURES
source
1..171
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Wassiljewskaja"
/db_xref="taxon:3702"
/clone="587A06"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
misc_feature
1..171
/note="T-DNA flanking sequence
left border"

ORIGIN
Query Match 43.4%; Score 38.2; DB 9; Length 171;
Best Local Similarity 55.2%; Pred. No. 0.0018;
Matches 37; Conservative 12; Mismatches 18; Indels 0; Gaps 0;

QY 7 GGUUCCGUAGUGUAGUGUUAUCGUCGUAACACGCGAAAGUCCCGGUCGAA 66
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 16 GGTTCGTGCTGTAGTGTGTTATCATCGTCAGTCTAACACACTCCGGTTCGAA 75

QY 67 ACCGGGC 73
Db 76 CCGGGGC 82

RESULT 4
LOCUS
DEFINITION
PUDG26TD 2M_0.6_1.0_KB Zea mays genomic clone ZMBFai179F03,
genomic survey sequence.
CC004711
ACCESSION
VERSION
CC004711.1
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays

23
UGGUUAUCAGUUCGCUACACGCGAAAGUCCCGGUCGAAACCGGGCACUACAAA 82
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 78 TGGTTATCAGTCAGTCTACACACTCAAGGTCCTCCGGTTCGAACCGGGCGCGCAA 19

QY 83 ACC 85
Db 18 ATC 16

RESULT 3

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

## REFERENCE

## AUTHORS

1. (bases 1 to 174)  
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Uterback, T.,  
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J., and  
Bennetzen, J.

## TITLE

Maize Genomics Consortium

## JOURNAL

Unpublished (2003)

## COMMENT

Contact: Cathy Whitelaw

## TIGR

7712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

## FEATURES

source

1. 174

Location/Qualifiers

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone\_lib="ZMMBTal79F03"

/clone\_lib="ZM 0.6 1.0 KB"

/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high

Cot selected genomic DNA library"

## ORIGIN

Query Match 37.5%; Score 33; DB 8; Length 174;  
Best Local Similarity 57.5%; Pred. No. 0.17; Indels 0; Gaps 0;  
Matches 42; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

QY 14 GUAGUGUAGUGGUUAUCAGUUCGCTTAACACGCGAAAGGUCGCCGUGCGAAACCGGC 73  
DB 66 GTAGTTCAGATGTAGACGCTCGCTTAGCAGCGAGCGCGGATCGATACCCCGC 125  
QY 74 ACUACAAAACCA 86  
DB 126 ATCTCAAAATCA 138

## RESULT 5

## BX176513

## LOCUS

BX176513 177 bp DNA linear GSS 13-MAR-2003  
Danio rerio genomic clone DKEY-176E1, genomic survey sequence.

## ACCESSION

## BX176513

## VERSION

## BX176513.1

## KEYWORDS

## GSS.

## SOURCE

## Danio rerio

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.

1. (bases 1 to 177)

Humphray, S.J., Huckle, E. and Durham, J.L.

Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome

Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquerry@sanger.ac.uk Unpublished

This sequence was generated from the T7 end of BAC 176E1. 176E1 is

part of the Daniokey BAC Library created by R. Plasterk and N.V.

Keygens. Further details:

http://www.sanger.ac.uk/Projects/D\_rerio/.

Location/Qualifiers

1. 177

/organism="Danio rerio"

/mol\_type="genomic DNA"

/db\_xref="taxon:7955"

/clone="DKEY-176E1"

/tissue type="Testis"

/note="Vector pIndigoBAC-536"

## ORIGIN

Query Match 35.0%; Score 30.8; DB 7; Length 126;  
Best Local Similarity 52.7%; Pred. No. 1.1;  
Matches 39; Conservative 8; Mismatches 27; Indels 0; Gaps 0;

QY 15 UAGUGUAGUGGUUAUCAGUUCGCCUAAACACGCGAAAGGUCGCCGUGCGAAACCGGCA 74  
DB 51 TAGTCTAGTGTAGAGCGGCTTGCTTAGAAGCGCAAGCCCTGGGTTCGTCCTCCAGCT 110  
QY 75 CUACAAAACCAAC 88  
DB 111 CTGAAAAAAGAAC 124

RESULT 7

BX203515

LOCUS

DEFINITION

Accession

Version

Keywords

Query Match 35.9%; Score 31.6; DB 9; Length 177;  
Best Local Similarity 54.8%; Pred. No. 0.57;  
Matches 34; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

QY 12 CCGUAGUGUAGUGGUUAUCAGUUCGCCUAAACACGCGAAAGGUCGCCGUGCGAAACCGG 71  
DB 115 CAGTAATGACGTGGCAGCACGTTTCGCTCACAGCAAGAGGTGCTGTTGAGCCCTTG 174  
QY 72 GC 73  
DB 175 GC 176

RESULT 6

H04713

LOCUS

DEFINITION

Accession

Version

Keywords

Source

Organism

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1. (bases 1 to 126)

Schwartz, P.E., Grieshaber, N.A., Grieshaber, S.S. and Majack, R.A.

An expressed sequence tag from in vitro embryonic rat vascular

smooth muscle cells

Unpublished (1995)

Contact: Phillip E. Schwartz

Pediatrics and Cellular and Structural Biology

University of Colorado Health Sciences Center

4200 East Ninth Avenue, Denver, Colorado, 80262, USA

Tel: (303) 270-4569

Fax: (303) 270-8353

Email: schwartz\_p@defiance.hsc.colorado.edu.

Location/Qualifiers

1. 126

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/strain="Sprague-Dawley"

/db\_xref="taxon:10116"

/clone="e17-1.3.3"

/clone\_lib="Lambda Zap Express library of P. E. Schwartz

(9/93)"

/note="developmental-stage=Embryonic Day 17

Post-fertilization; tissue-type=Aorta; cell-type=Vascular

smooth muscle cell; sex=Male."

ORIGIN

Query Match 35.0%; Score 30.8; DB 7; Length 126;  
Best Local Similarity 52.7%; Pred. No. 1.1;  
Matches 39; Conservative 8; Mismatches 27; Indels 0; Gaps 0;

QY 15 UAGUGUAGUGGUUAUCAGUUCGCCUAAACACGCGAAAGGUCGCCGUGCGAAACCGGCA 74  
DB 51 TAGTCTAGTGTAGAGCGGCTTGCTTAGAAGCGCAAGCCCTGGGTTCGTCCTCCAGCT 110  
QY 75 CUACAAAACCAAC 88  
DB 111 CTGAAAAAAGAAC 124

RESULT 7

BX203515

LOCUS

DEFINITION

Accession

Version

Keywords





follows: 5' adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor sequence:

5'-ATTAGAGCGGCGGCGGCACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.08 kb (range .73-1.37 kb). 13/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA).Corp."

## ORIGIN

Query Match 33.4%; Score 29.4; DB 6; Length 117;  
Best Local Similarity 57.1%; Pred. No. 3.6;  
Matches 36; Conservative 6; Mismatches 21; Indels 0; Gaps 0;  
QY 25 GUUAUCAUGUCCUACACGCGAAAGUCCCGUUCGAAACCGCGCACUACAAAAC 84  
Db 28 GGTAGAGCGTCTGCTAGCAGCGCGCTGGTTCGCTCTCAGCTCTACAAAAA 87  
QY 85 CAA 87  
Db 88 AAA 90

## RESULT 10

BI845249/c

## LOCUS

DEFINITION 130 bp mRNA linear EST 04-OCT-2001  
Is92a07.y1 Zebrafish neuronal Danio rerio cDNA clone IMAGE:5077140  
5', mRNA sequence.

## ACCESSION

BI845249

## VERSION

BI845249.1

## KEYWORDS

EST

## SOURCE

Danio rerio (zebrafish)

## ORGANISM

Danio rerio

## REFERENCE

AUTHORS

Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Matra, M.,

Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,

Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,

Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,

Ritter, E., Kohn, S., Shin, T., Jackson, F., Cardenas, M., McCann, R.,

Waterston, R. and Wilson, R.

WashU Zebrafish EST Project 1998

Unpublished (1998)

Contact: Stephen L. Johnson

Washington University School

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrafish@watson.wustl.edu

cDNA Library constructed by S. Lin DNA Sequencing by: Washington

University Genome Sequencing Center Clone distribution: the

I.M.A.G.E. Consortium/UNL, send email to: info@image.llnl.gov

Seq primer: 73 ET from Amersham.

Location/Qualifiers

1..130

/organism="Danio rerio"

/mol\_type="mRNA"

/db\_xref="taxon:7955"

/clones="IMAGE:5077140"

/sex="mixed"

/dev\_stages="adult"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="Zebrafish neuronal"

/note="Organ: brain; Vector: pBluescript (modified);

Site 1: DraIII(Y); Site 2: DraIII(X); Library is cloned

directionally between the DraIII(X) and DraIII(Y) sites

and has been amplified. Library constructed by S. Lin."

## ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

Length

Pred. No.

4.4;

33.4%;

Score

29.2;

DB

4;

Length

130;

53.4%;

Pred. No.

4.4;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

4.5;

33.2%;

Score

29.2;

DB

8;

Length

166;

52.7%;

Pred. No.

SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 166)  
REFERENCE  
AUTHORS  
Zhao,S., Niernan,W., Feldblyum,T., Malek,J., Shatsman,S.,  
Akinret,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de  
Jong,P. and Fraser,C.M.  
Mouse BAC End Sequences from Library RPCI-23  
Unpublished (1999)  
Other\_GSSs: RPCI-23-476P9.TV  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-23. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)  
or from Resea ch Genetics (info@resgen.com). BAC end page:  
http://www.tigr.org/tdb/bac\_ends/mouse/bac\_end\_intro.html  
Plate: 476 row: F column: 9  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source  
1..166  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-23-476P9"  
/sex="Female"  
/lab\_host="DH10B"  
/clone\_lib="RPCI-23"  
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:  
EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or  
brain genomic DNA was isolated and partially digested  
with a combination of EcoRI and EcoRI Methylase. Size  
selected DNA was cloned into the pBACe3.6 vector at the  
EcoRI sites. The ligation products were transformed into  
DH10B electrocompetent cells (BRL Life Technologies). "

ORIGIN  
Query Match 33.0%; Score 29; DB 8; Length 166;  
Best Local Similarity 52.2%; Pred. No. 5.4;  
Matches 36; Conservative 8; Mismatches 25; Indels 0; Gaps 0;  
QY 15 UAGUGAGUGGUAUCACGUAAGGUCGCCGUAACACCGGAAGGUCGCCGUAACACCGGCA 74  
DB 89 TAGTTTCAGTTGATGATGCTCTCTTACATGACATAGCCCTGGGTTTGATCCTCAGAA 148  
QY 75 CUACAAAAA 83  
DB 149 CAACATAA 157

RESULT 14  
BX244513 175 bp DNA linear GSS 13-MAR-2003  
LOCUS  
Danio rerio genomic clone DKEY-233N3, genomic survey sequence.  
DEFINITION  
BX244513  
ACCESSION  
BX244513.1 GI:28166847  
VERSION  
GSS.  
KEYWORDS  
SOURCE  
ORGANISM  
Danio rerio (zebrafish)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;  
Cypriniformes; Cyprinidae; Danio.  
1 (bases 1 to 175)  
REFERENCE  
AUTHORS  
Humphray,S.J., Huckle,E. and Durham,J.L.

LOCUS  
RPCI-23-153F14.TJ RPCI-23 Mus musculus genomic clone  
DEFINITION  
RPCI-23-153F14, genomic survey sequence.  
ACCESSION  
AZ261468 1 GI:9469876  
VERSION  
GSS.  
KEYWORDS  
SOURCE  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 157)  
REFERENCE  
AUTHORS  
Zhao,S., Niernan,W., Feldblyum,T., Malek,J., Shatsman,S.,  
Akinret,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de  
Jong,P. and Fraser,C.M.  
Mouse BAC End Sequences from Library RPCI-23  
Unpublished (1999)  
Other\_GSSs: RPCI-23-153F14.TV  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-23. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)  
or from Resea ch Genetics (info@resgen.com). BAC end page:  
http://www.tigr.org/tdb/bac\_ends/mouse/bac\_end\_intro.html  
Plate: 153 row: F column: 14  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source  
1..157  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-23-153F14"  
/sex="Female"  
/lab\_host="DH10B"  
/clone\_lib="RPCI-23"  
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:  
EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or  
brain genomic DNA was isolated and partially digested  
with a combination of EcoRI and EcoRI Methylase. Size  
selected DNA was cloned into the pBACe3.6 vector at the  
EcoRI sites. The ligation products were transformed into  
DH10B electrocompetent cells (BRL Life Technologies). "

ORIGIN  
Query Match 33.0%; Score 29; DB 8; Length 157;  
Best Local Similarity 54.1%; Pred. No. 5.4;  
Matches 33; Conservative 8; Mismatches 20; Indels 0; Gaps 0;  
QY 26 UUAUCACGUCGUAACACGGAAGGUCGCCGUAACACCGGCAACAAACC 85  
DB 132 TTAAATAATTGCTAGCATGACAAAGCCCTGGGTTCAATCTTAGCCTACATAACC 73  
QY 86 A 86  
DB 72 A 72

RESULT 13  
AZ118437 166 bp DNA linear GSS 12-MAY-2000  
LOCUS  
RPCI-23-476P9.TJ RPCI-23 Mus musculus genomic clone RPCI-23-476P9,  
genomic survey sequence.  
DEFINITION  
AZ118437  
ACCESSION  
AZ118437.1 GI:7782376  
VERSION  
GSS.  
KEYWORDS

**TITLE** Direct Submission  
**JOURNAL** Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquiesanger.ac.uk Unpublished  
**COMMENT** This sequence was generated from the T7 end of BAC 233N3. 233N3 is part of the Daniokey BAC Library created by R. Piasterik and N.V. Keygene. Further details: [http://www.sanger.ac.uk/Projects/D\\_rerio/](http://www.sanger.ac.uk/Projects/D_rerio/).  
**FEATURES** Location/Qualifiers  
 source 1..175  
 /organism="Danio rerio"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7955"  
 /clone="DKEY-233N3"  
 /tissue\_type="Testis"  
 /note="vector pindigoBAC-536"

**ORIGIN**  
 Query Match 33.0%; Score 29; DB 9; Length 175;  
 Best Local Similarity 50.7%; Pred. No. 5.4;  
 Matches 35; Conservative 9; Mismatches 25; Indels 0; Gaps 0;  
 QY 12 CCUAGUGUAGUGUUAUCAGUUCGCUAACACGCGGAAAGGUCGCCGCUUCCGAAACCGG 71  
 Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 42 CCATGCTCAGTGTAGTACTGCGCATCACAGAGAGGTCCTGTTAAGTCCCG 101  
 QY 72 GCACUACAA 80  
 Db |||:|||||  
 102 GCTGGACGA 110

**RESULT 15**  
**LOCUS** C0739878  
**DEFINITION** Spermophilus lateralis (golden-mantled ground squirrel) clone 25e03 5', mRNA sequence.  
**ACCESSION** C0739878  
**VERSION** C0739878.1 GI:50827148  
**KEYWORDS** EST.  
**SOURCE** Spermophilus lateralis (golden-mantled ground squirrel)  
**ORGANISM** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae; Spermophilus.  
**REFERENCE** 1. (bases 1 to 158)  
**AUTHORS** Williams,D.R., Gracey,A.Y., Martin,S.L., Hughes,M.A., Li,W., Rogers,J. and Cossins,A.R.  
**TITLE** Microarray analysis of transcriptional changes during hibernation in the golden mantled ground squirrel, Spermophilus lateralis  
**JOURNAL** Unpublished (2004)  
**COMMENT** Contact: Andrew R. Cossins  
 Laboratory for Environmental Gene Regulation  
 University of Liverpool  
 School of Biological Sciences, The Biosciences Building, Crown Street, Liverpool, United Kingdom, L69 7ZB  
 Tel: +44(0)151-795-4510  
 Fax: +44(0)151-795-4431  
 Email: cossins@liv.ac.uk  
 Vector has been trimmed from this EST.  
 Plate: 25 row: e column: 03  
 Seq primer: TripleX 5'LD (5'-CTCGGGAAGCGGCCCATTTGTTGTTGT-3')  
 High quality sequence stop: 158.  
**FEATURES** Location/Qualifiers  
 source 1..158  
 /organism="Spermophilus lateralis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:76772"  
 /clone="25e03"  
 /sex="Male & female"  
 /tissue\_type="brain"  
 /dev\_stage="Adult"  
 /lab\_host="E.coli Electromax DH10B"  
 /clone\_lib="squirrel brain library 1"

**ORIGIN**  
 Query Match 32.7%; Score 28.8; DB 7; Length 158;  
 Best Local Similarity 56.2%; Pred. No. 6.4;  
 Matches 36; Conservative 6; Mismatches 22; Indels 0; Gaps 0;  
 QY 25 GUUAUCAGUUCGCUUACACGCGAAAGGUCGCCGCUUCCGAAACCGGGCACUACAAAAAC 84  
 Db |.:|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 16 GGTAGAGCGCTTGCTGTAGCATGCGTAAGGCCCTGGGTTGCTCTCAGCCCTGTAAAAA 75  
 QY 85 CAAC 88  
 Db |||||  
 76 CAAC 79

Search completed: October 31, 2004, 05:19:40  
 Job time : 1731 secs

**ORIGIN**  
 Query Match 32.7%; Score 28.8; DB 7; Length 158;  
 Best Local Similarity 56.2%; Pred. No. 6.4;  
 Matches 36; Conservative 6; Mismatches 22; Indels 0; Gaps 0;  
 QY 25 GUUAUCAGUUCGCUUACACGCGAAAGGUCGCCGCUUCCGAAACCGGGCACUACAAAAAC 84  
 Db |.:|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 16 GGTAGAGCGCTTGCTGTAGCATGCGTAAGGCCCTGGGTTGCTCTCAGCCCTGTAAAAA 75  
 QY 85 CAAC 88  
 Db |||||  
 76 CAAC 79

**ORIGIN**  
 Query Match 32.7%; Score 28.8; DB 7; Length 158;  
 Best Local Similarity 56.2%; Pred. No. 6.4;  
 Matches 36; Conservative 6; Mismatches 22; Indels 0; Gaps 0;  
 QY 25 GUUAUCAGUUCGCUUACACGCGAAAGGUCGCCGCUUCCGAAACCGGGCACUACAAAAAC 84  
 Db |.:|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
 16 GGTAGAGCGCTTGCTGTAGCATGCGTAAGGCCCTGGGTTGCTCTCAGCCCTGTAAAAA 75  
 QY 85 CAAC 88  
 Db |||||  
 76 CAAC 79



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 31, 2004, 01:51:44 ; Search time 852.61 Seconds  
(without alignments)  
1774.868 Million cell updates/sec

Title: US-09-974-974-1  
Perfect score: 32  
Sequence: 1 gguccggcgaugagagagucucuc 32

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 1924030

Minimum DB seq length: 0  
Maximum DB seq length: 60

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sv.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	32	100.0	32	6	AX453842	AX453842 Sequence
2	32	100.0	32	6	BD143498	BD143498 Nucleic a
3	21.4	66.9	33	6	AX453844	AX453844 Sequence
4	21.4	66.9	33	6	BD143500	BD143500 Nucleic a
5	17.8	55.6	60	6	CQ547198	CQ547198 Sequence
6	17.4	54.4	29	6	BD197186	BD197186 Method an
7	17.4	54.4	40	6	AX453853	AX453853 Sequence
8	17.4	54.4	47	6	AR291770	AR291770 Sequence
9	16.6	51.9	30	6	AR179707	AR179707 Sequence
10	16.6	51.9	30	6	BD222050	BD222050 SH2 domai
11	16.4	51.2	27	6	BD162072	BD162072 Method fo
12	16.4	51.2	55	6	AX211297	AX211297 Sequence
13	16.2	50.6	39	6	AR213624	AR213624 Sequence
14	16.2	50.6	39	6	BD057682	BD057682 Fusion pr
15	16.2	50.6	39	6	BD081512	BD081512 Soluble s
16	16.2	50.6	51	6	AR145871	AR145871 Sequence
17	16.2	50.6	60	6	CQ550477	CQ550477 Sequence
18	16	50.0	29	6	BD199490	BD199490 Method an
19	16	50.0	53	6	AR105779	AR105779 Sequence

C	20	15.8	49.4	38	6	AX218725	AX218725 Sequence
	21	15.6	48.8	50	6	AX161018	AX161018 Sequence
	22	15.6	48.8	60	6	CQ553925	CQ553925 Sequence
C	23	15.4	48.1	25	6	AX533469	AX533469 Sequence
	24	15.4	48.1	25	6	AX533472	AX533472 Sequence
	25	15.4	48.1	36	6	AR041983	AR041983 Sequence
	26	15.4	48.1	36	6	AX637441	AX637441 Sequence
	27	15.4	48.1	38	6	AR334837	AR334837 Sequence
C	28	15.4	48.1	51	6	AX190288	AX190288 Sequence
	29	15.4	48.1	57	4	BOVOST	M31318 Bovine oste
	30	15.4	48.1	60	6	CQ535421	CQ535421 Sequence
C	31	15.2	47.5	28	6	I33133	I33133 Sequence 3
	32	15.2	47.5	60	6	CQ551175	CQ551175 Sequence
C	33	15	46.9	25	6	AX533467	AX533467 Sequence
	34	15	46.9	25	6	AX533468	AX533468 Sequence
	35	15	46.9	27	6	BD206476	BD206476 Enzymatic
	36	15	46.9	29	6	BD202230	BD202230 Method an
C	37	15	46.9	50	6	CQ008519	CQ008519 Sequence
	38	15	46.9	51	6	CQ007060	CQ007060 Sequence
C	39	15	46.9	58	6	AX924083	AX924083 Sequence
	40	15	46.9	60	6	CQ538656	CQ538656 Sequence
C	41	14.8	46.2	34	6	I29828	I29828 Sequence 14
	42	14.8	46.2	38	6	AR331695	AR331695 Sequence
C	43	14.8	46.2	38	6	AR332379	AR332379 Sequence
	44	14.8	46.2	38	6	AR332820	AR332820 Sequence
C	45	14.8	46.2	38	6	AX222448	AX222448 Sequence

## ALIGNMENTS

RESULT 1  
AX453842  
LOCUS AX453842 32 bp RNA linear PAT 06-JUL-2002  
DEFINITION Sequence 1 from Patent EP1213351.  
ACCESSION AX453842  
VERSION AX453842.1 GI:21713511  
KEYWORDS synthetic construct  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS Taira,K., Warashina,M. and Warashina,T.  
TITLE Nucleic acid enzymes acquiring an activity for cleaving a target  
JOURNAL rna by recognising another molecule  
Patent: EP 1213351-A 1 12-JUN-2002;  
National Institute of Advanced Industrial Science and Technology (JP)

FEATURES  
Source Location/Qualifiers  
1..32  
/organism="synthetic construct"  
/mol\_type="unassigned RNA"  
/db\_xref="taxon:32630"  
/note="maxizyme-constituting RNA molecule"

ORIGIN

Query Match 100.0%; Score 32; DB 6; Length 32;  
Best Local Similarity 71.9%; Pred. No. 0.00063;  
Matches 23; Conservative 9; Mismatches 0; Indels. 0; Gaps 0;

Qy 1 GGUCGCGCCUGAGAGAGUGAGUCUCUC 32  
||:|||||:|||||:|||||:|||||:|||||:  
Db 1 GGTCTGGCTTGATGAGATGATGACTTTC 32

RESULT 2

BD143498  
LOCUS BD143498 32 bp RNA linear PAT 17-JAN-2003  
DEFINITION Nucleic acid enzyme acquiring activity of cleaving other specific  
target RNA by recognizing RNA molecule.  
ACCESSION BD143498  
VERSION BD143498.1 GI:27849256  
KEYWORDS JP 2002119283-A/1.



```
ORIGIN
SOURCE
  /mol_type="unassigned DNA"
  /db_xref="taxon:9606"
REFERENCE
  Query Match 55.6%; Score 17.8; DB 6; Length 60;
  Best Local Similarity 71.4%; Pred. No. 3.5e+03;
  Matches 15; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 1 GUCCUGGCCUGAGAGAGUG 21
  ||:|||||:|||||
Db 37 GGTCTGGCCTGAAGACAGTG 57

RESULT 6
BD197186
LOCUS
  29 bp RNA linear PAT 17-JUL-2003
DEFINITION
  Method and reagent for treating diseases or conditions concerning
  molecule participating in vasculogenic response.
ACCESSION
  BD197186
VERSION
  BD197186.1 GI:33006956
KEYWORDS
  JP 2002509721-A/212.
SOURCE
  synthetic construct
  artificial sequences.
ORGANISM
  1 (bases 1 to 29)
  Pavco,P.A., Roberts,E., Jarvis,T., Coeshott,C. and Mcswiggen,J.A.
  Method and reagent for treating diseases or conditions concerning
  molecule participating in vasculogenic response
  Patent: JP 2002509721-A 212 02-APR-2002;
JOURNAL
  RIBOZYME PHARMACEUTICALS INC
COMMENT
  OS Artificial Sequence
  PN JP 2002509721-A/212
  PD 02-APR-2002
  PF 24-MAR-1999 JP 2000541291
  PR 27-MAR-1998 US 60/079678
  PI PAMELA A PAVCO,ELISABETH ROBERTS,THALE JARVIS,CLAIRE COESHOTT,
  PI JAMES A MCSWIGGEN
  PC
  C12N15/09,A61K31/7088,A61K31/7125,A61K48/00,A61P3/10,A61P17/06,PC
  A61P29/00,
  PC A61P35/00,A61P43/00,C12N5/10,C12N9/00//A61K35/76,C12N15/00,PC
  C12N5/00
  CC Synthesized Hammerhead Ribozyme
  CC The letter 'n' stands for any base or bases forming a loop or
  CC that may contain multiple nucleic acid analogues or 2'-CC
  deoxynucleotides.
FH Key Location/Qualifiers
FT source 1..29
  /organism="synthetic construct"
  /mol_type="genomic RNA"
  /db_xref="taxon:32630"
FEATURES
  source
  1..29
  Location/Qualifiers
ORIGIN
  Query Match 54.4%; Score 17.4; DB 6; Length 29;
  Best Local Similarity 53.6%; Pred. No. 5.4e+03;
  Matches 15; Conservative 6; Mismatches 7; Indels 0; Gaps 0;
QY 3 UCCUGGCCUGAGAGAGUGAGUCUCU 30
  :||:|||||:|||||
Db 2 TCTTGGCCTGATGAGNCGAAGAGTTCT 29

RESULT 7
AX453853
LOCUS
  40 bp RNA linear PAT 06-JUL-2002
DEFINITION
  Sequence 12 from Patent EP1213351.
ACCESSION
  AX453853
VERSION
  AX453853.1 GI:21713522
KEYWORDS
  source

SOURCE
  /mol_type="unassigned DNA"
  /db_xref="taxon:9606"
REFERENCE
  Query Match 55.6%; Score 17.8; DB 6; Length 60;
  Best Local Similarity 71.4%; Pred. No. 3.5e+03;
  Matches 15; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 1 GUCCUGGCCUGAGAGAGUG 21
  ||:|||||:|||||
Db 37 GGTCTGGCCTGAAGACAGTG 57

RESULT 8
AR291770/c
LOCUS
  47 bp DNA linear PAT 12-JUN-2003
DEFINITION
  Sequence 3505 from patent US 6537751.
ACCESSION
  AR291770
VERSION
  AR291770.1 GI:31679054
KEYWORDS
  Unknown.
SOURCE
  Unknown.
  Unclassified.
REFERENCE
  1 (bases 1 to 47)
  Cohen,D., Chumakov,I. and Blumenfeld,M.
  Biallelic markers for use in constructing a high density
  disequilibrium map of the human genome
  Patent: US 6537751-A 3505.25-MAR-2003;
JOURNAL
  Location/Qualifiers
FEATURES
  source
  1..47
  /organism="unknown"
  /mol_type="genomic DNA"
ORIGIN
  Query Match 54.4%; Score 17.4; DB 6; Length 47;
  Best Local Similarity 51.7%; Pred. No. 5.5e+03;
  Matches 15; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
QY 2 GUCCUGGCCUGAGAGAGUGAGUCUCU 30
  :||:|||||:|||||
Db 32 GTCCATGCTGATGATGCTCTGAGCTCT 4

RESULT 9
AR179707
LOCUS
  30 bp DNA linear PAT 20-APR-2002
DEFINITION
  Sequence 7 from patent US 6326482.
ACCESSION
  AR179707
VERSION
  AR179707.1 GI:20221262
KEYWORDS
  Unknown.
  Unclassified.
REFERENCE
  1 (bases 1 to 30)
  Stewart,T.A. and Lu,Y.
  SH2 domain-containing peptides
  Patent: US 6326482-A 7 04-DEC-2001;
JOURNAL
  Location/Qualifiers
FEATURES
  source
  1..30
  Location/Qualifiers
```

REFERENCE  
1 (bases 1 to 27)  
AUTHORS  
Tanaka,M. and Kikuchi,A.  
TITLE  
Method for detecting or quantitating protein  
JOURNAL  
Patent: JP 2002191364-A 8 09-JUL-2002;

RESULT 13			
AR213624	AR213624	39 bp	DNA
LOCUS	Sequence 58 from patent US 6405989.	linear	PAT 25-SEP-2002
DEFINITION			



```
ACCESSION AR213624
VERSION AR213624.1 GI:23310903
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 39)
AUTHORS Davis,M.E., White,R.A., Saunders,C., Polin,R., Kristiansen,K.,
Ballone,M. and Grossman,G.
TITLE Rollable sports base
JOURNAL Patent: US 6405989-A 58 18-JUN-2002;
FEATURES
    source
        1..39
        /organism="unknown"
        /mol_type="genomic DNA"
ORIGIN
Query Match 50.6%; Score 16.2; DB 6; Length 39;
Best Local Similarity 62.1%; Pred. No. 2e+04;
Matches 18; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 1 GGUCUCCGCGCTGCTGAGGGTGACGATCCC 29
    |||||:||||:|||||
Db 4 GGGCCCGGGCTGCTGAGGGTGACGATCCC 32
    |||||:||||:|||||
RESULT 14
BD057682
LOCUS BD057682
DEFINITION Fusion proteins comprising bacteriophage coat protein and a
single-chain T cell receptor.
ACCESSION BD057682
VERSION BD057682.1 GI:22603288
KEYWORDS
SOURCE Aspergillus tubingensis
ORGANISM Aspergillus tubingensis
REFERENCE 1 (bases 1 to 39)
AUTHORS Weidanz,J.A., Card,K.F. and Wong,H.C.
TITLE Fusion proteins comprising bacteriophage coat protein and a
single-chain T cell receptor
JOURNAL Patent: JP 2001514503-A 58 11-SEP-2001;
COMMENT SUNOL MOLECULAR CORP
PN JP 2001514503-A/58
PD 11-SEP-2001
PF 05-MAR-1998 JP 1998537984
PR 07-MAR-1997 US 08/813781
PI JON A WEIDANZ,KIMBERLYN F CARD,HING C WONG
PC C12Q1/68,C12N7/01,C12N15/70
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
    source
        1..39
        /organism="Aspergillus tubingensis"
        /mol_type="genomic DNA"
        /db_xref="taxon:5068"
ORIGIN
Query Match 50.6%; Score 16.2; DB 6; Length 39;
Best Local Similarity 62.1%; Pred. No. 2e+04;
Matches 18; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 1 GGUCUCCGCGCTGCTGAGGGTGACGATCCC 29
    |||||:||||:|||||
Db 4 GGGCCCGGGCTGCTGAGGGTGACGATCCC 32
    |||||:||||:|||||
RESULT 15
BD081512
LOCUS BD081512
DEFINITION Soluble single-chain T-cell receptor proteins.
ACCESSION BD081512
VERSION BD081512.1 GI:22627115
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 39)
AUTHORS Weidanz,J.A., Card,K.F. and Wong,H.C.
TITLE Soluble single-chain T-cell receptor proteins
JOURNAL Patent: JP 2001519143-A 58 23-OCT-2001;
COMMENT SUNOL MOLECULAR CORP
PN JP 2001519143-A/58
PD 23-OCT-2001
PF 28-SEP-1998 JP 2000514936
PR 02-OCT-1997 US 08/943086
PI JON A WEIDANZ,KIMBERLYN F CARD,HING C WONG
PC C12N15/09,A61K38/00,A61K39/395,A61P43/00,C07K14/725,C07K16/28,
C12P21/02//
PC C12P21/08,C12N15/00,A61K37/02
CC Description of Artificial Sequence: primer
FH Key Location/Qualifiers
FT source 1..39
FT /organism='Artificial Sequence'
FEATURES
    source
        1..39
        Location/Qualifiers
        1..39
        /organism="synthetic construct"
        /mol_type="genomic DNA"
        /db_xref="taxon:32630"
ORIGIN
Query Match 50.6%; Score 16.2; DB 6; Length 39;
Best Local Similarity 62.1%; Pred. No. 2e+04;
Matches 18; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
QY 1 GGUCUCCGCGCTGCTGAGGGTGACGATCCC 29
    |||||:||||:|||||
Db 4 GGGCCCGGGCTGCTGAGGGTGACGATCCC 32
    |||||:||||:|||||
Search completed: October 31, 2004, 03:24:20
Job time : 862.61 secs
```



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 31, 2004, 01:51:44 ; Search time 719.39 Seconds  
(without alignments)  
1774.868 Million cell updates/sec

Title: US-09-974-974-2  
 Perfect score: 27  
 Sequence: 1 GUGUGACUGUUGCAUCGAAACCGUCC 27

Scoring table: IDENTITY\_NUC  
Gapop 10.0 : Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 1924030

```
Minimum DB seq length: 0
Maximum DB seq length: 60
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:★

```

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %			ID	Description
		Match	Length	DB		
1	27	100.0	27	6	AX453843	AX453843 Sequence
2	27	100.0	27	6	BD143499	BD143499 Nucleic a
3	18	66.7	29	6	AX453845	AX453845 Sequence
4	18	66.7	29	6	BD143501	BD143501 Nucleic a
C 5	17	63.0	44	6	AR183095	AR183095 Sequence
C 6	16	59.3	49	6	CQ760641	CQ760641 Sequence
C 7	15.6	57.8	45	9	HSCD857116	250829 H.sapiens m
C 8	15.2	56.3	43	6	AX297640	AX297640 Sequence
9	15	55.6	24	6	AX445798	AX445798 Sequence
C 10	15	55.6	51	6	CQ007470	CQ007470 Sequence
11	14.8	54.8	32	6	AR090340	AR090340 Sequence
12	14.8	54.8	32	6	AR197375	AR197375 Sequence
13	14.8	54.8	32	6	AR259529	AR259529 Sequence
C 14	14.8	54.8	33	6	BD274953	BD274953 POLYNUCLE
C 15	14.8	54.8	33	6	AR474082	AR474082 Sequence
C 16	14.6	54.1	22	6	AR078172	AR078172 Sequence
17	14.6	54.1	28	6	AX351504	AX351504 Sequence
C 18	14.6	54.1	54	10	AF224227	AF224227 Mus muscu
19	14.4	53.3	58	6	AX773271	AX773271 Sequence

## ALIGNMENTS

RESULT 1  
AX453843                    27 bp      RNA          linear      PAT 06-JUL-2002

LOCUS  
DEFINITION Sequence 2 from Patent EP1213351.  
ACCESSION AX453843  
VERSION AX453843.1 GI:21713512

KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
         artificial sequences.

REFERENCE  
1 Taira,K., Warashina,M. and Warashina,T.  
TITLE Nucleic acid enzymes acquiring an activity for cleaving a target  
rna by recognising another molecule  
JOURNAL Patent: EP 1213351-A 2 12-JUN-2002;  
National Institute of Advanced Industrial Science and Technology  
(JP)

FEATURES  
source Location/Qualifiers  
1..27  
         /organism="synthetic construct"  
         /mol\_type="unassigned RNA"  
         /db\_xref="taxon:32630"  
         /note="maxizyme-constituting RNA molecule"

ORIGIN

Query Match                 100.0%; Score 27; DB 6; Length 27;  
Best Local Similarity 74.1; Pred.No. 0.013;  
Matches 20; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GUCUGAGTGUCAUCGAAACCGGUCC 27  
| | | | | | | | | | | | | | | | | |  
Db 1 GTCTGACTGTTCATCGAACC GGSTCC 27

RESULT 2  
BD143499                    27 bp      RNA          linear      PAT 17-JAN-2003

LOCUS  
DEFINITION Nucleic acid enzyme acquiring activity of cleaving other specific  
target RNA by recognizing RNA molecule.  
ACCESSION BD143499  
VERSION BD143499.1 GI:27849257  
KEYWORDS JP 2002119283-A/2.

Db	12	TTCATCGAATACCGGTCC	29
RESULT 4			
BD143501			
LOCUS	BD143501	29 bp	linear
DEFINITION	Nucleic acid enzyme acquiring activity of cleaving other specific target RNA by recognizing RNA molecule.		
ACCESSION	BD143501		PAT 17-JAN-2003
VERSION	BD143501.1	GI:27849259	
VERSION	BD143501.1	GI:27849259	

KEYWORDS	JP 2002119283-A/4.
SOURCE	synthetic construct
ORGANISM	synthetic construct
	artificial sequences.
REFERENCE	1 (bases 1 to 29)
AUTHORS	Taira, K., Warashina, M. and Warashina, T.
TITLE	Nucleic acid enzyme acquiring activity of cleaving other specific target RNA by recognizing RNA molecule
JOURNAL	Patent: JP 2002119283-A 4 23-APR-2002; DIRECTOR GENERAL OF NATIONAL INSTITUTE OF ADVANCED INDUSTRIAL SCIENCE AND TECHNOLOGY, JAPAN BIO INDUSTRY ASSOCIATION,
COMMENT	OS Artificial Sequence PN JP 2002119283-A/4 PD 23-APR-2002 PF 13-OCT-2000 JP 2000313320

PC C12N15/09, A61K9/127, A61K38/46, A61K48/00, A61P31/12, A61P35/00, PC  
C12N1/15,  
PC C12N1/19, C12N1/21, C12N5/10, C12N9/00, C12Q1/25, C12Q1/68, C12N15/  
PC C00 A61K37/54

PC	00, A6, A3, 7, 54,
PC	C12N5/00
CC	Description of Artificial Sequence: maxizyme-constituting RNA
CC	molecule
FH	Key
FH	Location/Qualifiers
FT	source
FT	1. .29

```

FEATURES
  source
    Location/Qualifiers
      FT
        /organism="Artificial sequence"
        1..29
        /organism="synthetic construct"

```

```

ORIGIN
Query Match          66.7%;   Score 18;   DB 6;   Length 29;
Best Local Similarity 77.8%;   Pred. No. 4.9e+02;
Matches 14;   Conservative 4;   Mismatches 0;   Indels 0;   Gaps 0;
QY 10 UTCAUCGAAACCGGUCC 27
    :|:|:|:|:|:|:|:|
Db 12 TTCATCGAAACCGGTCC 29

```

RESULT 5	AR183095/c	AR183095	44 bp	DNA	linear	PAT 20-APR-2002
LOCUS		Sequence	72	from patent	US 6340461.	
DEFINITION		AR183095				
ACCESSION		AR183095				
VERSION		AR183095	1	GT:202226588		

VERSION	AKI83095.1	Gr:20226688
KEYWORDS		
SOURCE	Unknown.	
ORGANISM	Unknown.	
REFERENCE	Unclassified.	
AUTHORS	1 (bases 1 to 44)	
	Terman, D. Stephen.	

JOURNAL  
FEATURES  
source  
1. .44  
Location/Qualifiers  
Patent: US 6340461-A 72 22-JAN-2002;  
diseases

```

ORIGIN
/organism="unknown"
/mol_type="unassigned DNA"

```

```
Query Match      63.0%; Score 17; DB 6; Length 44;
Best Local Similarity 60.0%; Pred. No. 1.7e+03;
Matches 15; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 CUGACUGUUCGAAACCGGGUCC 27
    |||::||: |||::||
Db 25 CAGACTGTTTCATAGGATCGGATCC 1

RESULT 6
LOCUS CQ760641/c 49 bp DNA linear PAT 03-MAR-2004
DEFINITION Sequence 83 from Patent WO2004003229.
ACCESSION CQ760641
VERSION CQ760641.1 GI:44904144
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Nex,B.R., Vogel,U., Rockenbauer,E. and Bukowy,Z.K.
TITLE Disease risk estimating method using sequence polymorphisms in a
JOURNAL specific region of chromosome 19
Patent: WO 2004003229-A 83 08-JAN-2004;
Aarhus University (DK) ; Arbejdsmilj Institutttet (National
Institute of Occupational Health) (DK)
FEATURES
source Location/Qualifiers
1..49
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Probe"
ORIGIN
Query Match      59.3%; Score 16; DB 6; Length 49;
Best Local Similarity 54.2%; Pred. No. 5.5e+03;
Matches 13; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 GUGUGACUGUUCGAAACCGGG 24
    |||::||: |||::||
Db 28 GTCGAGTATTGATCGAACCCAGG 5

RESULT 7
HSCD85716/c
LOCUS HSCD85716 45 bp mRNA linear PRI 15-MAR-1996
DEFINITION H.sapiens mRNA for T cell receptor (ID CD85716).
ACCESSION Z50829
VERSION Z50829.1 GI:1004252
KEYWORDS constant region; joining region; T cell receptor; variable region.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 45)
Wang,E.C., Moss,P.A., Frodsham,P., Lehner,P.J., Bell,J.I. and
Borysiewicz,L.K.
CD8highCD57+ T lymphocytes in normal, healthy individuals are
oligoclonal and respond to human cytomegalovirus
JOURNAL J. Immunol. 155 (10), 5046-5056 (1995)
MEDLINE 96062329
PUBMED 7594513
REFERENCE 2
AUTHORS Wang,E.C., Moss,P.A., Frodsham,P., Lehner,P.J., Bell,J.I. and
Borysiewicz,L.K.
TITLE Oligoclonal CD8+, CD57+ T cells in normal, healthy individuals and
their responses to HCMV
JOURNAL Unpublished
3 (bases 1 to 45)
Wang,E.C.
AUTHORS Direct Submission
TITLE Submitted (17-AUG-1995) Eddie CY Wang, Medicine, University of
JOURNAL
```

```
Wales of College of, Medicine, Heath Park, Cardiff, South
Glamorgan, CF4 4XN, Wales UK
Location/Qualifiers
1..45
/organism="Homo sapiens"
/mol_type="mRNA"
/isolate="SG"
/db_xref="taxon:9606"
1..12
/note="end T cell receptor V beta 3"
V_segment 13..27
C_region /note="CDR3"
28..45
J_segment /note="beginning J beta 2.1"
ORIGIN
Query Match      57.8%; Score 15.6; DB 9; Length 45;
Best Local Similarity 53.6%; Pred. No. 8.7e+03;
Matches 14; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 ACUGUUCGAAACCGGGUCC 27
    |||::||: |||::||
Db 40 ACTGCTCATTGTACCGAGTCC 19

RESULT 8
LOCUS AX297640/c 43 bp DNA linear PAT 21-NOV-2001.
DEFINITION Sequence 9402 from Patent WO0179548.
ACCESSION AX297640
VERSION AX297640.1 GI:17059331
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Barany,F., Zirvi,M., Gerry,N.P., Favis,R. and Kliman,R.
TITLE Method of designing addressable array for detection of nucleic acid
JOURNAL sequence differences using ligase detection reaction
Patent: WO 0179548-A 9402 25-OCT-2001;
CORNELL RESEARCH FOUNDATION, INC. (US)
FEATURES
source Location/Qualifiers
1..43
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Hypothetical Probe Sequence"
ORIGIN
Query Match      56.3%; Score 15.2; DB 6; Length 43;
Best Local Similarity 65.0%; Pred. No. 1.4e+04;
Matches 13; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 7 CUGUUCGAAACCGGGUCC 26
    |||::||: |||::||
Db 35 CTGTCCATCGAATCGGGTC 16

RESULT 9
LOCUS AX445798 24 bp DNA linear PAT 03-JUL-2002
DEFINITION Sequence 2253 from Patent WO0216649.
ACCESSION AX445798
VERSION AX445798.1 GI:21694697
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Gunderson,K.
TITLE Probes and decoder oligonucleotides
JOURNAL Patent: WO 0216649-A 2253 28-FEB-2002;
Illumina, Inc. (US)
```

```
FEATURES
  source
    Location/Qualifiers
      1..24
        /organism="synthetic construct"
        /mol_type="unassigned DNA"
        /db_xref="taxon:32630"
        /note="Computer Generated Probe Sequence."
ORIGIN
  Query Match 55.6%; Score 15; DB 6; Length 24;
  Best Local Similarity 60.9%; Pred. No. 1.6e+04;
  Matches 14; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 3 CUGACUGUUCUACGAAACCGGGU 25
  ||||| : : : |||||
Db 1 CTGACTGCATCTCGAAGCGGGT 23

RESULT 10
LOCUS CQ007470/c 51 bp DNA linear PAT 16-JAN-2004
DEFINITION Sequence 6110 from Patent WO0147944.
ACCESSION CQ007470
VERSION CQ007470.1 GI:41014119
KEYWORDS Homo sapiens (human)
SOURCE
  ORGANISM
    Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
    1 Shimkets,R.A. and Leach,M.
    Nucleic acids containing single nucleotide polymorphisms and
    methods of use thereof
    JOURNAL Patent: WO 0147944-A 6110 05-JUL-2001;
    Curagen Corporation (US)
  FEATURES
    source
      1..51
        /organism="Homo sapiens"
        /mol_type="unassigned DNA"
        /db_xref="taxon:9606"
        /note="Accession number cg43988954"
ORIGIN
  Query Match 55.6%; Score 15; DB 6; Length 51;
  Best Local Similarity 60.9%; Pred. No. 1.8e+04;
  Matches 14; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 GUCUGACUGUUCUACGAAACCGG 23
  ||| : : : |||||
Db 23 GCCTGGATGCTCTTCGAAACGG 1

RESULT 11
LOCUS AR090340 32 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 460 from patent US 5994076.
ACCESSION AR090340
VERSION AR090340.1 GI:10017095
KEYWORDS
  SOURCE
    Unknown.
  ORGANISM
    Unclassified.
  REFERENCE
    1 (bases 1 to 32)
    Chenchik,A., Jokhadze,G. and Bibilashvili,R.
    Methods of assaying differential expression
    JOURNAL Patent: US 5994076-A 460 30-NOV-1999;
    Location/Qualifiers
  FEATURES
    source
      1..32
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
  Query Match 54.8%; Score 14.8; DB 6; Length 32;
  Best Local Similarity 53.8%; Pred. No. 2.1e+04;
  Matches 14; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 2 UCUGACUGUUCUACGAAACCGGUCC 27
  ||| : : : |||||
Db 6 TCAGATCTCCATTCAAACCGAGTCC 31

RESULT 12
LOCUS AR197375 32 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 460 from patent US 6352829.
ACCESSION AR197375
VERSION AR197375.1 GI:20247224
KEYWORDS
  SOURCE
    Unknown.
  ORGANISM
    Unclassified.
  REFERENCE
    1 (bases 1 to 32)
    Chenchik,A., Jokhadze,G. and Bibilashvili,R.
    Methods of assaying differential expression
    JOURNAL Patent: US 6352829-A 460 05-MAR-2002;
    Location/Qualifiers
  FEATURES
    source
      1..32
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
  Query Match 54.8%; Score 14.8; DB 6; Length 32;
  Best Local Similarity 53.8%; Pred. No. 2.1e+04;
  Matches 14; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 2 UCUGACUGUUCUACGAAACCGGUCC 27
  ||| : : : |||||
Db 6 TCAGATCTCCATTCAAACCGAGTCC 31

RESULT 13
LOCUS AR259529 32 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 460 from patent US 6489455.
ACCESSION AR259529
VERSION AR259529.1 GI:27310040
KEYWORDS
  SOURCE
    Unknown.
  ORGANISM
    Unclassified.
  REFERENCE
    1 (bases 1 to 32)
    Chenchik,A., Jokhadze,G. and Bibilashvili,R.
    Methods of assaying differential expression
    JOURNAL Patent: US 6489455-A 460 03-DEC-2002;
    Location/Qualifiers
  FEATURES
    source
      1..32
        /organism="unknown"
        /mol_type="genomic DNA"
ORIGIN
  Query Match 54.8%; Score 14.8; DB 6; Length 32;
  Best Local Similarity 53.8%; Pred. No. 2.1e+04;
  Matches 14; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 2 UCUGACUGUUCUACGAAACCGGUCC 27
  ||| : : : |||||
Db 6 TCAGATCTCCATTCAAACCGAGTCC 31

RESULT 14
LOCUS BD274953 33 bp DNA linear PAT 17-JUL-2003
DEFINITION POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY.
ACCESSION BD274953
VERSION BD274953.1 GI:33084721
KEYWORDS JP 2002539786-A/17;
  SOURCE
    synthetic construct
```

ORGANISM synthetic construct  
artificial sequences.  
REFERENCE 1 (bases 1 to 33)  
AUTHORS Shimkets,R.A.  
TITLE POLYNUCLEOTIDES AND PROTEINS ENCODED THEREBY  
JOURNAL Patent: JP 2002538786-A 17 19-NOV-2002;  
Curagen Corporation,Richard A Shimkets  
COMMENT OS Artificial Sequence  
PN JP 2002538786-A/17  
PD 19-NOV-2002  
PF 09-MAR-2000 JP 2000603363  
PR 08-MAR-2000 US 09/520781,09-MAR-1999 US 60/123667 PI  
richard a shimkets  
CC Description of Artificial Sequence: Primer  
FH Key Location/Qualifiers  
FEATURES source  
1..33  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"  
ORIGIN  
Query Match 54.8%; Score 14.8; DB 6; Length 33;  
Best Local Similarity 53.8%; Pred. No. 2.1e+04;  
Matches 14; Conservative 5; Mismatches 7; Indels 0; Gaps 0;  
QY 2 UCUGACUGUUCGAAACCGGUCC 27  
Db 26 TCAGAACTCTCTGGGAAACCGGATCC 1  
RESULT 15  
AR474082/c  
LOCUS AR474082 33 bp DNA linear PAT 20-FEB-2004  
DEFINITION Sequence 33 from patent US 6689866.  
ACCESSION AR474082  
VERSION AR474082.1 GI:42712839  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 33)  
AUTHORS Shimkets,R.A.  
TITLE Polynucleotides and proteins encoded thereby  
JOURNAL Patent: US 6689866-A 33 10-FEB-2004;  
FEATURES source  
1..33  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 54.8%; Score 14.8; DB 6; Length 33;  
Best Local Similarity 53.8%; Pred. No. 2.1e+04;  
Matches 14; Conservative 5; Mismatches 7; Indels 0; Gaps 0;  
QY 2 UCUGACUGUUCGAAACCGGUCC 27  
Db 26 TCAGAACTCTCTGGGAAACCGGATCC 1

Search completed: October 31, 2004, 03:24:23  
Job time : 722.39 secs

